DUKE MATHEMATICAL JOURNAL

Volume 5, Number 3
SEPTEMBER, 1939

DUKE UNIVERSITY PRESS
DURHAM, N. C.

DUKE MATHEMATICAL JOURNAL

This periodical is published quarterly under the auspices of Duke University by Duke University Press at Durham, North Carolina. It is printed at Mt. Royal and Guilford Avenues, Baltimore, Maryland, by the Waverly Press.

Entered as second class matter at the Post Office, Durham, North Carolina. Additional entry at the Post Office, Baltimore, Maryland.

The subscription price for the current volume is four dollars, postpaid; back volumes, five dollars each, carriage extra; single numbers, one dollar and twenty-five cents, postpaid. Individual members of the Mathematical Association of America may subscribe at half price; mention of such membership should be made when subscribing. Subscriptions, orders for back numbers, and notice of change of address should be sent to Duke University Press, Durham, North Carolina.

Manuscripts and editorial correspondence should be addressed to Duke Mathematical Journal, 4785 Duke Station, Durham, North Carolina. An Author's Manual containing detailed information about the preparation of papers for publication will be sent on request.

Manuscripts submitted for consideration by the editors should be type-written, with the exception of the formulas, which may be filled in by hand. It is recommended, though not required, that footnotes be numbered consecutively with Arabic numerals, and that they be placed at the end of the manuscript. A footnote will appear, however, at the bottom of the printed page of the journal on which reference is made to it. The author should keep a duplicate copy of the manuscript as a precaution against loss in transit and as an aid in reading proof. The original manuscript will not accompany the galley proof sent to the author unless requested. Such a request may delay the appearance of the article. Drawings must be made in black India ink on Bristol board and fully lettered.

Authors are entitled to one hundred free reprints. Additional copies will be furnished at cost. All reprints will be furnished with covers unless the contrary is specifically requested.

Made in United States of America

WAVERLY PRESS, INC.
BALTIMORE, U. S. A.

T
com
theo
para
of s
poin
curv
This
form
P^2_n
P'^2_n
pair
set a
by
P^2_n
this
grou
grue
is o

W
sion
"reg
"iso
Huc

In
proj
of P
prop
so a
set
$S_{k(n}$

P:

The
min

R
at it
Univ

CONFIGURATIONS DEFINED BY THETA FUNCTIONS

By Arthur B. Coble

The aspects of theta function theory which I wish to present today have come to my own attention in connection with an effort to develop an invariant theory for the Cremona group in a projective space which would in some sense parallel the invariant theory of the projective group in that space. One type of such a theory arises in the plane in connection with the idea of a set of n points P^2_n as the carrier of the totality Σ^2_n of all the complete linear systems of curves which can be defined by their multiplicities at the points of the set. This duality between P^2_n and Σ^2_n is obviously invariant under projective transformation, i.e., if P^2_n is carried into P'^2_n by a projectivity π, then Σ^2_n defined by P^2_n is carried by π into Σ'^2_n defined by P'^2_n. But it is also true that if P^2_n and P'^2_n are such that $n - \rho$ pairs p_i, p'_i drawn from the two sets are corresponding pairs of a Cremona transformation τ for which the remaining ρ points in each set are the direct and inverse F-points of τ, then also Σ^2_n defined by P^2_n is carried by τ into Σ'^2_n defined by P'^2_n. Under these circumstances we say that the set P^2_n is *congruent* to the set P'^2_n under the Cremona transformation τ and regard this notion of congruence of sets of points P^2_n as the extension to the Cremona group of the notion of projectivity of sets of points. In this notion of congruence, as in projectivity, the *order* of the points in the related sets P^2_n, P'^2_n is obviously material.

When we extend this notion to sets of points P^k_n in spaces of higher dimension k, it is necessary to confine the Cremona transformations to elements of the "regular" Cremona group, i.e., to transformations which can be defined by "isolated" F-points—transformations which are termed "punctual" by Miss Hudson.

In developing this notion of congruence, we find it convenient to eliminate projectivity by using the obvious canonical form in which the first $k + 2$ points of P^k_n are taken to be the reference points, and the unit point and the factors of proportionality in the coördinates of the remaining $n - k - 2$ points are so adjusted that the last coördinate of each is the same, say u. Then the set P^k_n is uniquely determined by the coördinates of a point P in a space $S_{k(n-k-2)}$,

$$P: \quad x_{ij}:x_{i'j'}: \cdots :u \qquad (i, i', \cdots = k + 3, \cdots, n; j, j', \cdots = 1, \cdots, k).$$

The ratios $x_{ij}:u$ are then absolute projective invariants or double ratios determined by P^k_n. Naturally this mapping of sets P^k_n in S_k on points P of

Received June 8, 1939; an invited address before the American Mathematical Society at its meeting held April 7–8, 1939, in conjunction with the centennial celebration of Duke University.

$S_{k(n-k-2)}$ has many "singular" or F-loci corresponding to "special" sets P_n^k, but generically it is birational.

A first inquiry with respect to this mapping concerns the variation of P induced by a change in the *order* of the points of P_n^k. This change is represented by a Cremona group $G_{n!}$ in $S_{k(n-k-2)}$, the $n!$ points P in a conjugate set of $G_{n!}$ representing the $n!$ orders in a P_n^k which corresponds to one of the $n!$ points P. The invariant theory of this $G_{n!}$ is essentially the projective invariant theory of P_n^k regarded as a *symmetric* set of n points in S_k. It is thus one rather immediate extension of the theory of invariants of a single binary form.

A second and more important inquiry with respect to this mapping concerns the variation of P induced by the transition from P_n^k to a congruent $P_n'^k$. If we use the theorem that the regular Cremona group in S_k is generated by collineations and the single regular transformation of type $x_i x_i' = 1$ ($i = 1, \cdots, k + 1$), then it is necessary to add to $G_{n!}$ only one additional generator to obtain the new type of variation. This additional generator has the very simple form $x_{ij} x_{ij}' = 1$, $uu' = 1$. With $G_{n!}$ it generates a Cremona group $G_{n,k}$ in $S_{k(n-k-2)}$ which is, except in a few instances to be noticed later, infinite and discontinuous. A conjugate set of points P under $G_{n,k}$ represents the aggregate of sets P_n^k congruent in some order to one ordered set in the aggregate.

We may say that the group $G_{n,k}$ in $S_{k(n-k-2)}$, a group without absolute constants, represents the capacity of the regular Cremona transformations in S_k with no more than n F-points to distort the projective properties of sets of no more than n points.

The nature of the group $G_{n,k}$ may be studied with comparative ease through its isomorphism with a linear group. In the case $G_{n,2}$ let the set of points P_n^2 be on a cuspidal cubic C^3, $x_1 = t^3$, $x_2 = t$, $x_3 = 1$, with parameters $t_1, \cdots, t_n$. The cusp is $t = \infty$, the flex is $t = 0$, and the collinear condition is $s_1 + s_2 + s_3 = 0$. Under the quadratic transformation A_{123} with F-points at t_1, t_2, t_3 the cubic C^3 is transformed into a cuspidal cubic C'^3, the point $p(t)$ of C^3 going into the point $p'(t)$ of C'^3. On C'^3, however, the collinear condition is $s_1 + s_2 + s_3 + t_1 + t_2 + t_3 = 0$. If then on C'^3 we make the change of parameter $t' = t + \frac{1}{3}(t_1 + t_2 + t_3)$ so as to restore the *form* $s_1' + s_2' + s_3' = 0$ of the collinear condition, the cubic C'^3 can be projected back upon C^3 in such a way that $p'(t')$ falls on $p(t')$. But the inverse F-points on C'^3 of A_{123} are those whose parameters t are respectively $-(t_2 + t_3)$, $-(t_3 + t_1)$, $-(t_1 + t_2)$. Hence the linear transformation

$$L_{123}: \quad \begin{aligned} t_i' &= t_i + \tfrac{1}{3}(t_1 + t_2 + t_3) - (t_1 + t_2 + t_3) \qquad &(i = 1, 2, 3),\\ t_j' &= t_j + \tfrac{1}{3}(t_1 + t_2 + t_3) \qquad &(j = 4, \cdots, n) \end{aligned}$$

furnishes that set of points $t_1', \cdots, t_n'$ on C^3 which is congruent to the original set of points $t_1, \cdots, t_n$ on C^3 under A_{123}. The reordering of P_n^2 being obviously equivalent to a permutation of $t_1, \cdots, t_n$, we see that the group $G_{n,2}$ is simply isomorphic with the collineation group on $t_1, \cdots, t_n$ generated by the permutations of $t_1, \cdots, t_n$ and the element L_{123}. We say the collineation group rather than the linear group since the transformation $t' = \rho t$ represents a collineation which carries C^3 into itself and P_n^2 into a projectively equivalent $P_n'^2$.

For the set P_n^k, by using a cuspidal C^{k+1} in S_k we find similarly a collineation group, isomorphic with $G_{n,k}$ which is generated by the permutations of $t_1, \cdots, t_n$ and the element

$$L_{1,\cdots,k+1}: \begin{array}{ll} t_i' = t_i + \dfrac{k-1}{k+1}(t_1 + \cdots + t_{k+1}) - (t_1 + \cdots + t_{k+1}) & (i = 1, \cdots, k+1), \\ t_j' = t_j + \dfrac{k-1}{k+1}(t_1 + \cdots + t_{k+1}) & (j = k+2, \cdots, n). \end{array}$$

In a few particular cases congruence implies projectivity so that the isomorphism between the collineation group and $G_{n,k}$ is 1 to 2. It is easy to verify that this linear group has a quadratic invariant form

$$\{n(k-1) - (k+1)^2\}\{(\textstyle\sum t^2)\} - (k-1)\{(\textstyle\sum t)^2\}.$$

If this form is *definite* and *negative*, the linear group, and therefore also the isomorphic $G_{n,k}$, is finite. The condition for this is $n(k-1) - (k+1)^2 < 0$, or, in general, $n \leqq k + 3$, except that for $k = 2, 3, 4$ this upper limit is increased by 3, 1, 1, respectively.

When the Cremona group $G_{n,k}$ in $\Sigma_{k(n-k-2)}$ is finite, it has an invariant theory of type quite similar to that of a finite projective group. When, however, it is infinite and discontinuous, the situation is not so simple. The group then may or may not have *invariants*, an invariant being a primal in Σ which is transformed into itself by all the elements of the group to within extraneous factors which arise from the "principal loci" of the particular transformation employed. However, the group has, in common with infinite continuous Cremona groups, the property of having invariant linear systems of primals. For example, in the first infinite cases which arise in S_2 and S_3 in connection with the point sets P_9^2 and P_8^3, the sets are on a unique elliptic curve which has projective invariants S, T and a linear system of invariants $S^3 + kT^2$. This incidence of point set and normal elliptic curve being invariant under regular Cremona transformation with F-points at P_9^2, P_8^3, respectively, the corresponding infinite Cremona groups $G_{9,2}$ and $G_{8,3}$ in Σ_{10} and Σ_9, respectively, have *pencils* of invariant primals. In these, and in the few analogous simple instances, there exists also a discontinuous aggregate of invariant primals of increasing orders. Thus for P_9^2 the condition that P_9^2 be the set of r-fold points of a curve of order $3r$ ($r \geqq 2$) yields for each value of r an invariant primal of $G_{9,2}$. Similarly, the conditions that P_{10}^2 be on a cubic curve, or that P_{10}^3 be on a quadric surface, yield invariant primals of $G_{10,2}$ and $G_{10,3}$, respectively.

For larger values of n the conditions on P_n^k that the points lie on a normal elliptic curve E^{k+1} in S_k yield a linear system with base E of primals in $\Sigma_{k(n-k-2)}$ invariant under $G_{n,k}$, the base of this system being the locus of points P which map sets P_n^k of this special character. Similarly, for $k \geqq 3$, the conditions on P_n^k that the n-points lie on a normal rational surface R^{k-1} in S_k yield a linear system with base R of primals invariant under $G_{n,k}$. It is not difficult to see that the base E is contained in the base R though the linear systems of primals

defined by E and R have different orders. There has not thus far been observed in S_k ($k \geqq 4$) types of three- or more-dimensional manifolds which are unaltered in type by regular Cremona transformation in S_k, so that the question as to the existence of other invariant linear systems of $G_{n,k}$ remains open.

The isomorphism of the group $G_{n,k}$ with a linear group has already been mentioned. This linear group was represented on n variables, but it is usually more convenient to represent it on $n + 1$ variables which may be identified with respectively the order and the multiplicity of a linear system of primals in S_k at the n points of P_n^k. The group then has also an invariant linear form, but, to offset this, it has only integral coefficients. Any group of this character has for each value of the integer ν an invariant subgroup of finite index which consists of those elements which reduce to the identity mod ν. The finite factor group of this invariant subgroup may be represented by the elements of the linear group with coefficients reduced mod ν. Thus the infinite Cremona groups $G_{n,k}$ have a sequence of invariant subgroups and a corresponding sequence of *finite* factor groups $G_{n,k}(\nu)$. In the finite cases these invariant subgroups are usually merely the identity, and the factor groups therefore the finite $G_{n,k}$ itself.

In the first infinite $G_{n,k} = G_{9,2}$ these invariant subgroups determined by ν are characterized by the fact that in $\Sigma_{k(n-k-2)}$ there exists a primal $M(\nu)$ whose points are invariant under the operations of the corresponding invariant subgroup. Then the $G_{9,2}$ effects upon the points of the invariant $M(\nu)$ only those permutations of a group isomorphic with the factor group $G_{9,2}(\nu)$. This primal $M(\nu)$ is the locus of points P whose corresponding sets P_9^2 have the property of being the 9 ν-fold points of a curve of order 3ν. In further cases these invariant subgroups are not so easily defined in terms of P_n^k itself.

Of particular interest in such a connection as has just been described is the first prime modulus 2. For, the linear group has a quadratic invariant, and the polarized quadratic may be regarded as a null system [since $x_1y_2 + x_2y_1 \equiv x_1y_2 - x_2y_1 \pmod 2$]. In the theory of the integral linear transformation of the $2p$ periods of the theta functions of p variables a similar invariant null system appears, and under such transformation the $2^{2p} - 1$ proper half periods are permuted under the finite linear group with coefficients reduced mod 2 which has this invariant null system. Thus it is natural to expect that the factor group $G_{n,k}(2)$ is either this theta factor group or a subgroup of it.

A detailed examination of $G_{n,k}(2)$ shows that this is indeed true. According as $n, k \equiv 0, 1, 2, 3 \pmod 4$ we find 16 cases. In two of these cases, $G_{n,k}(2)$ is simply isomorphic to the group of the above null-system, say $G_p(2)$, a simple group. In four other cases, $G_{n,k}(2)$ is simply isomorphic to that subgroup, also simple, of $G_p(2)$ which leaves one of the 2^{2p} odd or even theta functions unaltered. In the remaining ten cases $G_{n,k}(2)$ contains an invariant subgroup, usually Abelian, whose order is a power of 2, and the factor group of $G_{n,k}(2)$ with respect to this invariant subgroup is again either $G_p(2)$ itself, or the subgroup of $G_p(2)$ which leaves an odd or even theta unaltered.

With the composition of $G_{n,k}(2)$ thus determined there remains to be con-

sidered the subgroup ${}_{(2)}G_{n,k}$ of $G_{n,k}$ whose factor group is this $G_{n,k}(2)$. In the case ${}_{(2)}G_{9,2}$ this group is Abelian, of additive character, with parameters of the form $\lambda_1 v_1 + \cdots + \lambda_8 v_8$, where $\lambda_1, \cdots, \lambda_8$ are positive integers and $v_1, \cdots, v_8$ general complex numbers. In the next case of $G_{10,2}$ the group ${}_{(2)}G_{10,2}$ may be described as isomorphic to the group of Cremona transformations which has an invariant rational sextic. If t is a parameter on such a sextic, it is therefore isomorphic to an infinite discontinuous group with elements of type $t' = (at + b)/(ct + d)$, and this isomorphism is simple if there exist no Cremona transformations T for which every point of the sextic is fixed. In a recent number of the Roma Rendiconti, Pompili gave a construction for such a transformation T, which, however, turns out to be erroneous.

In developing the theta functions of p complex variables, we customarily express the variables in terms of the $2p$ periods by means of a set of $2p$ *real* parameters $g_1, \cdots, g_p, g_1', \cdots, g_p'$ so that a value system u has a *period characteristic* $u = \begin{bmatrix} g \\ g' \end{bmatrix}$. The 2^{2p} half periods, including the zero half period, are then given by $\begin{bmatrix} \epsilon \\ \epsilon' \end{bmatrix}_2$, the $2p$ values ϵ, ϵ' being 0 or 1 with a denominator 2. The various theta functions are also distinguished by the values of a *theta characteristic* $\begin{Bmatrix} h \\ h' \end{Bmatrix}$, h, h' also real, in such a way that

$$\vartheta \begin{Bmatrix} h \\ h' \end{Bmatrix} \left(u + \begin{bmatrix} g \\ g' \end{bmatrix} \right) = \vartheta \begin{Bmatrix} h + g \\ h' + g' \end{Bmatrix} (u) \cdot E,$$

where E is a properly chosen exponential factor. In particular the 2^{2p} odd and even theta functions are given by $\vartheta \begin{Bmatrix} \eta \\ \eta' \end{Bmatrix}_2 (u)$, where η, η' are 0 or 1 with denominator 2. It is in terms of this "characteristic notation" that the majority of the theta formulas have been given.

However, in order to make the transition from the group $G_{n,k}(2)$ attached to sets of points P_n^k to the theta modular group a quite different notation, the so-called "basis notation" is practically essential. In this notation we use $2p + 2$ indices $1, 2, \cdots, 2p + 2$ and name the 2^{2p} half periods by an even number of indices $P_{1,2,\cdots,2k} = P_{2k+1,\cdots,2p+2}$, the zero half period being given by $P = P_{1,\cdots,2p+2}$. On the other hand the 2^{2p} odd and even theta functions are named by sets of $p + 1 \pm 2k$ indices as in

$$\vartheta_{1,2,\cdots,p+1-2k}(u) = \vartheta_{p+1-2k+1,\cdots,2p+2}(u).$$

These functions are even or odd with k. All the tactical relations among the half periods and functions are now quite elementary. Thus $P_{i_1\cdots i_{2k}} + P_{j_1\cdots j_{2l}} = P_{i_1\cdots i_{2k} j_1 \cdots j_{2l}}$ with, as always, like indices canceling. Also

$$\vartheta_{i_1\cdots i_{p+1-2k}}(u + P_{j_1\cdots j_{2l}}) = E \cdot \vartheta_{i_1\cdots i_{p+1-2k} j_1 \cdots j_{2l}}(u).$$

Two half periods are syzygetic or azygetic according as they have or have not an even number of indices in common. A function does not, or does, vanish

for a half period according as k is even or odd in the number $p + 1 - 2k$ of their non-common indices.

The transition from the one notation to the other depends on the choice of the basis half periods $P_{i_1 i_2}$. A convenient choice is the following

$$\begin{bmatrix} 1 & 1 \cdots 1 \\ 1 & 1 \cdots 1 \end{bmatrix} = \begin{bmatrix} P_{12} & P_{34} & P_{56} & \cdots P_{2p-1,2p} \\ P_{2p+2,1} & P_{2p+2,1,2,3} & P_{2p+2,1,\cdots,5} & \cdots P_{2p,2p+1} \end{bmatrix}.$$

For this choice the function

$$\vartheta(u) = \vartheta \begin{Bmatrix} 0 & 0 \cdots 0 \\ 0 & 0 \cdots 0 \end{Bmatrix} (u)$$

is

$$\vartheta_{1,3,5,\cdots,2p+1}(u) = \vartheta_{2,4,6,\cdots,2p+2}(u).$$

For this choice, if the indices are divided into complementary sets $a_1, \cdots, a_{p+2}, b_1, \cdots, b_p$, we have the following typical relations among the even theta squares:

$$p = 1: \qquad \sum_{i=1}^{3} (-1)^{a_i} \vartheta^2_{a_i b_1} \vartheta^2_{a_i b_1}(u) \equiv 0,$$

$$p = 2: \qquad \sum_{i=1}^{4} (-1)^{a_i} \vartheta^2_{a_i b_1 b_2} \vartheta^2_{a_i b_1 b_2}(u) \equiv 0,$$

$$p = 3: \qquad \sum_{i=1}^{5} (-1)^{a_i} \vartheta^2_{a_i b_1 b_2 b_3} \vartheta^2_{a_i b_1 b_2 b_3}(u) - (-1)^{a} \vartheta^2 \vartheta^2(u) \equiv 0,$$

$$a = \sum_i a_i.$$

All of the respectively 3-, 4-, 5-termed relations among the theta squares can be obtained from one of these by the repeated use of two operations A and T, A being the addition of a half period and T being an integral linear transformation of the periods.

For $p = 4$ there are two relations connecting the 12 even theta squares formed like the above. These are

$$p = 4: \quad \begin{aligned} &\mathrm{I}: \sum_{i=1}^{6} (-1)^{a_i} \vartheta^2_{a_i b_1 \cdots b_4} \vartheta^2_{a_i b_1 \cdots b_4}(u) - \sum_{i=1}^{6} (-1)^{a+a_i} \vartheta^2_{a_i} \vartheta^2_{a_i}(u) \equiv 0, \\ &\mathrm{II}: \sum_{i=1}^{6} (-1)^{a_i} \epsilon_{a_i b_1 \cdots b_4} [\vartheta^2_{a_i b_1 \cdots b_4} \vartheta^2_{a_i}(u) + \vartheta^2_{a_i} \vartheta^2_{a_i b_1 \cdots b_4}(u)] \equiv 0, \end{aligned}$$

where $\epsilon_{a_i b_1 \cdots b_4}$ is the sign of the permutation $a_i b_1 \cdots b_4$ from the natural order of these indices. From these for $u = 0$ we have the modular relations

$$\mathrm{III}: \qquad \sum_{i=1}^{6} (-1)^{a_i} [\vartheta^4_{a_i b_1 \cdots b_4} - (-1)^{a} \vartheta^4_{a_i}] = 0,$$

$$\mathrm{IV}: \qquad \sum_{i=1}^{6} (-1)^{a_i} \epsilon_{a_i b_1 \cdots b_4} \vartheta^2_{a_i b_1 \cdots b_4} \vartheta^2_{a_i} = 0.$$

The last of these enables us to prove the interesting identity connecting the squares of the *ten* even theta functions which define the basis configuration, namely,

$$\text{V:} \qquad \sum_{i=1}^{10} (-1)^i \vartheta_i^2 \vartheta_i^2(u) \equiv 0,$$

whence

$$\text{VI:} \qquad \sum_{i=1}^{10} (-1)^i \vartheta_i^4 = 0.$$

If we multiply I by $(-1)^a$ and replace the second sum from V, we get

$$\text{VII:} \qquad \sum_{i=1}^{6} (-1)^{a+a_i} \vartheta^2_{a_ib_1\cdots b_4} \vartheta^2_{a_ib_1\cdots b_4}(u) + \sum_{j=1}^{4} (-1)^{b_j} \vartheta^2_{b_j} \vartheta^2_{b_j}(u) \equiv 0,$$

$$\text{VIII:} \qquad \sum_{i=1}^{6} (-1)^{a+a_i} \vartheta^4_{a_ib_1\cdots b_4} + \sum_{j=1}^{4} (-1)^{b_j} \vartheta^4_{b_j} = 0.$$

However, VII is merely that identity which arises from V by a transformation T which converts the given basis into a new basis.

Under the modular group M of integral linear transformations of the periods, the odd and even thetas undergo the permutations of a finite group $M(2)$ which is generated in a very simple way by a conjugate set of involutions. A particular generating involution $I_{i_1\cdots i_{2k}}$ is attached to a particular half period $P_{i_1\cdots i_{2k}}$. This involution leaves $P_{j_1\cdots j_{2l}}$ unaltered or transforms it into $P_{i_1\cdots i_{2k}j_1\cdots j_{2l}}$ according as $P_{j_1\cdots j_{2l}}$ is syzygetic or azygetic with $P_{i_1\cdots i_{2k}}$. This involution leaves $\vartheta_{j_1\cdots j_{p+1-2k}}$ unaltered or transforms it into $\vartheta_{j_1\cdots j_{p+1-2k}i_1\cdots i_{2k}}$ according as this new ϑ has not or has the same parity as the original ϑ.

The immediate foundation for the isomorphism between the group $G_{n,k}(2)$ associated with the Cremona group $G_{n,k}$ and the above modular group $M(2)$ is that $G_{n,k}$ and its isomorphic linear group $L_{n,k}$ are also generated by a conjugate set of involutions. Included in this set are *first* the transpositions of two variables (t_it_j) and *second* $L_{1,\ldots,k+1}$. If $n \geqq k+2$, these involutions are in the same conjugate set. When $L_{n,k}$ is infinite, this conjugate set is also infinite but it becomes finite when reduced mod 2. It becomes relatively easy to follow the behavior of these involutions under mutual transformation by observing that (t_it_j) and $L_{1,\ldots,k+1}$ are defined respectively by the linear forms $t_i - t_j$ and $t_1 + \cdots + t_{k+1}$. In establishing the isomorphism, the transpositions (t_it_j) are always identified with the involutions I_{ij} $(i, j = 1, \cdots, n)$, whereas $L_{1,\ldots,k+1}$ is identified with $I_{1,\ldots,k+1,i_1\cdots i_r}$, the $i_1 \cdots i_r$ being such properly chosen new indices with $n + r = 2p + 2$ that the theta involutions and the linear involutions combine in the same fashion. The case $G_{8,2}$ is a relatively simple illustration. We begin with

$$(t_it_j) : t_i - t_j \qquad : I_{ij} \qquad\qquad (i, j = 1, \cdots, 8),$$

$$L_{123} \ : t_1 + t_2 + t_3 : I_{1230},$$

the index 0 being added to make the number of indices of I_{1230} even, whence another index 9 is to be expected. On continuing the transformation process, we find that

$$L_{123456} = L_{123}L_{456}L_{123} \quad : t_1 + \cdots + t_6 \qquad : I_{123456} = I_{7890}\,,$$
$$L_{1,2\cdots 8} = L_{178}L_{123456}L_{178} : 2t_1 + t_2 + \cdots + t_8 : I_{19}\,.$$

It is now easy to verify that L_{123} and the transpositions transform these linear forms precisely like I_{1230} and I_{ij} transform the corresponding half periods. This establishes the isomorphism. The theta involutions I are those attached to all the half periods for which $\vartheta_0(u)$ vanishes *whence* $G_{8,2}(2)$, *or its finite isomorphic* $G_{8,2}$, *is isomorphic with that subgroup of* $M(2)$ *for* $p = 4$ *which leaves an even theta function* $\vartheta_0(u)$ *unaltered.*

In all of the finite cases of $G_{n,k}$ this abstract isomorphism between $G_{n,k}(2)$ and $M(2)$ is vitalized by an actual algebraic connection between the set of points $P_{n,k}$ and the theta functions concerned. Moreover, the effect of regular Cremona transformation upon P_n^k, and the effect of integral linear transformation of the periods upon the theta functions, are precisely those which are foreshadowed by the abstract isomorphism.

However, a necessary condition for this algebraic connection is the existence of a system of *three term* relations among the theta or theta modular functions. For, the set P_n^k projectively defines certain pencils of primes with $k - 1$ of its points as a base. Three primes of such a pencil, each on one of the remaining points, are linearly related. This relation must therefore appear in the theta functions which define the P_n^k.

For $p = 1$ and $p = 2$ these three-term relations are fairly numerous, and in the latter case, are interpretable in terms of a P_6^3 and the Weddle surface determined by it. In the case $p = 3$ with indices $1, \cdots, 8$ let

$$\vartheta_{1234}(0) = \vartheta_{5678}(0) = C_{1234} = C_{5678}\,,$$

and denote certain products of these C's as follows:

$$C_{ijk}^{(78)} = C_{ijk7}C_{ijk8} \qquad (i, j, k = 1, \cdots, 6),$$
$$C_{ij}^{(56)(78)} = C_{ij57}C_{ij58}C_{ij67}C_{ij68}\,.$$

Then one notable modular relation takes the form

$$\text{(R)} \qquad \sum_{i=1}^{3} \pm C_{i4}^{(56)(78)} = 0.$$

In this three-term relation the negative ratio of two terms is a double ratio D. If we think of the index 8 as isolated, D is a double ratio in a set P_7^2 of the four lines from p_7 to $p_1, \cdots, p_4$. If 8 is not isolated, D is a double ratio in a set P_8^3, the eight base points of a net of quadrics, of the four planes on either the line p_5p_6, or the line p_7p_8, to the points $p_1, \cdots, p_4$. Another type of

three-term relation involves the initial terms of the odd thetas ($p = 3$) with the variables u restricted to the Riemannian case. It yields an irrational equation of the quartic envelope with 7 of its 28 nodes at P_7^2. A third type of three-term relation for $p = 3$ arises from a four-term relation by imposing the requirement $\vartheta(u) = 0$ on the variables u_1, u_2, u_3. Then those of the four-term relations which contain $\vartheta(u)$ become three-term relations. The latter define a set P_7^3 in space and as u varies subject to $\vartheta(u) = 0$ a variable point p_8 is obtained which with P_7^3 makes up a P_8^3 which is the nodal set of an 8-nodal quartic surface. A fourth type for $p = 3$ occurs when $C = \vartheta(0) = 0$, and the functions are hyperelliptic. Then the three-term relations define a P_8^5 and a Weddle 3-way attached to it. Thus even in the relatively simple case of $p = 3$ the three-term relations (with the exception of (R) above) are known to exist only for special values of the moduli or the arguments. In the case $p = 4$ even this one exception seems to disappear.

The relation (R) has, however, some significance for the case $p = 4$ which appears as follows. If the period cell of the u's is mapped in 2-1 fashion ($\pm u$) by means of the theta squares upon a Kummer manifold K_p in S_{2^p-1}, the K_p is invariant under an Abelian collineation group $g_{2^{2p}}$ whose involutorial elements are defined by addition of the half periods, $u' = u + P_{i_1 \cdots i_{2k}}$. This $g_{2^{2p}}$ can be represented very simply by transformations with integral coefficients and thus is independent of the moduli of the functions. Given then the $g_{2^{2p}}$, and one of the 2^{2p} singular points of K_p which are determined by the half periods, say that one, O, determined by $u = 0$, the K_p itself is determined. As the $\frac{1}{2}p(p + 1)$ moduli change, this singular point configuration runs over a modular manifold M of dimension $\frac{1}{2}p(p + 1)$ in S_{2^p-1}. For $p = 1$, and $p = 2$, $\frac{1}{2}p(p + 1) = 2^p - 1$, and the modular manifold M covers the space of the $g_{2^{2p}}$. But for $p = 3$, and given g_{2^6} in S_7, M has the dimension 6, whence not every point of S_7 can be a singular point of a K_3 with the given g_{2^6}. If we write the relation (R) in the form

$$\sum_{i=1}^{3} \pm \sqrt{(C_{i4}^{(56)(78)})^2} = 0 \tag{R$_1$}$$

and express the theta squares in terms of the coördinates of g_{2^6}, the rationalized form of (R) is the equation of M.

An involution in $g_{2^{2p}}$, say $u' = u + P_{\alpha\beta}$, has two skew fixed spaces, F_1, F_2 of dimension S_{2^q-1} ($q = p - 1$). Each of these cuts K_p in 2^{2q} points, these being the points on K_p defined by quarter periods Q for which $2Q = P_{\alpha\beta}$. The elements of $g_{2^{2p}}$ which leave an F unaltered effect within it a $g_{2^{2q}}$. Thus the quarter period points on a space F of K_p behave like the half period points of a K_q ($q = p - 1$). Now it may be proved that if the half period points on K_q are defined by the zero values of the theta squares, $C_{i_1 \cdots i_{q+1-2k}}^2$, then the quarter period points of K_p in the F defined by $P_{\alpha\beta}$ are similarly defined by the theta products $C_{i_1 \cdots i_{q+1-2k}\alpha} C_{i_1 \cdots i_{q+1-2k}\beta}$ or, in the notation just used by $C_{i_1 \cdots i_{q+1-2k}}^{(\alpha\beta)}$.

Now Schottky has proved that the condition that the general theta functions of genus 4 with 10 moduli be Abelian with $3p - 3 = 9$ moduli is

$$(\mathrm{R}_2) \qquad \sum_{i=1}^{3} \pm \sqrt{C_{i4}^{(56)(78)(90)}} = 0.$$

Comparing this with the above form of (R_1), we see that Schottky's condition may be stated as follows:

The condition that the theta functions of genus 4 be Abelian is that one quarter period configuration shall be a half period configuration for functions of genus three.

For the general functions in four variables no three-term relations have been exhibited. For the Abelian functions, however, the set of modular relations formed as in (R_2) has been utilized by Schottky to define the sets of 9 and 10 nodes of a symmetroid. In connection then with these sets P_9^3 and P_{10}^3 and their associated sets P_9^4 and P_{10}^5 we have the desired algebraic relation indicated by the isomorphisms between $G_{n,k}(2)$ and the modular group $M(2)$, even though the $G_{n,k}$'s themselves are infinite.

When the functions are Abelian, another type of three-term relation ($p = 4$) occurs as a special case of a four-term relation because of the further modular condition $C_0 = 0$. It appears then that, if u_{ijk} is the linear term in the development of the odd Riemannian theta $\vartheta_{ijk}(u)$, we will have a linear identity connecting the three root-functions

$$\sqrt{u_{i_1j9}}\,\sqrt{u_{i_190}}, \quad \sqrt{u_{i_2j9}}\,\sqrt{u_{i_290}}, \quad \sqrt{u_{i_3j9}}\,\sqrt{u_{i_390}} \qquad (i_1, i_2, i_3, j = 1, \cdots, 8).$$

If then we set

$$F_{\alpha\beta} = \sqrt{u_{\alpha\beta9}}\,\sqrt{u_{\alpha90}}\,\sqrt{u_{\beta90}} \qquad (\alpha, \beta = 1, \cdots, 8; \alpha \neq \beta),$$

it is clear that three $F_{\alpha\beta}$ with a common index are linearly related and that all of the $F_{\alpha\beta}$ can be linearly expressed in terms of three like F_{12}, F_{13}, F_{23}. Hence the $F_{\alpha\beta}$'s can be regarded as the lines joining the points p_α, p_β of a planar P_8^2. As the u's vary the point x in the plane runs over the locus of the 9-th node of sextics with eight given nodes at P_8^2.

The only types of three-term relations which are known for generic p are those which exist in connection with the hyperelliptic functions for general values of the arguments. These serve to define the remaining sets P_n^k for which $G_{n,k}$ is finite, and the Weddle p-ways associated with them. These sets are more specifically P_{2p+2}^{2p-1} and P_{2p+1}^{2p-2}.

In every case in which the set P_n^k appears in algebraic connection with the theta functions it is true that the variation in the set P_n^k obtained by Cremona transformation to congruent sets, and the variation in the set obtained by integral period transformation, are identical.

UNIVERSITY OF ILLINOIS.

INVARIANTS

BY HERMANN WEYL

The *theory of invariants* came into existence about the middle of the nineteenth century somewhat like Minerva: a grown-up virgin, mailed in the shining armor of algebra, she sprang forth from Cayley's Jovian head. Her Athens over which she ruled and which she served as a tutelary and beneficent goddess was *projective geometry*. From the beginning she was dedicated to the proposition that all projective coördinate systems are created equal. Indeed, at that time the viewpoint of projective invariance was the one universally accepted in geometry. The rise of projective geometry had first been brought about by truly geometric stimuli, the study of conic sections, the theory of perspective and by the development of descriptive geometry, and the so-called synthetic direction of Steiner and von Staudt has confirmed the fertility of the projective attitude with respect to pure geometry.

However, its gaining such immense preponderance was, if I am not mistaken, due to algebraic rather than geometric reasons: namely, to the fact that the group of projectivities is expressed by the simplest of all continuous groups, the group of all homogeneous linear transformations. Plücker in the preface of his first work (*Analytisch-geometrische Entwicklungen*, vol. 1, 1828) openly espoused the ascendency of algebra, or, as he said, analysis, over geometry. So that perhaps one had better speak of geometric algebra than of algebraic geometry, namely, of an algebra which, in establishing its theorems and in the search for the proofs thereof, uses geometric terms and is guided by geometric intuition. The modern evolution, as far as it does not point its needle toward topology, has on the whole been marked by a trend of algebraization, notwithstanding the undeniable merits of the great school of Italian geometers.

The dictatorial rule of the projective idea in geometry was first successfully broken by the German astronomer and geometer Möbius. One is forced to realize that the group of all homogeneous linear transformations is not the only one worthy of consideration and capable of serving as the group of automorphisms in a geometric space. Möbius does not yet possess the general idea of a group; however, his notion of *Verwandtschaft* meets the same purpose in each special case he considers. The universal group theoretic interpretation was first

Received April 20, 1939; an invited address before the American Mathematical Society at its meeting held April 7–8, 1939 in conjunction with the centennial celebration of Duke University.

No explicit references to the literature were given in the address; they can readily be supplied from the author's book *The Classical Groups, their Invariants and Representations*, Princeton, 1939. For the general foundations of the theory of invariants compare in particular v.d. Waerden, Mathematische Annalen, vol. 113(1936), pp. 14–35.

promulgated in plain words by Felix Klein in his famous *Erlanger Programm* 1872, which is the classical document of the democratic platform in geometry yielding equal rights to each and every imaginable group. The adaptation of his standpoint to the study of invariants has been somewhat slow. Before I discuss the main problems of the theory of invariants I find it convenient to rephrase Klein's fundamental idea in slightly modernized and hence slightly more abstract terms.

One wishes to associate with the points P of a space numerical (i.e., reproducible) symbols x as their *coördinates*. In general this is possible in a conceptual manner, without pointing out the individual points with my finger, only with respect to a *frame of reference*, e.g., in Euclidean geometry with respect to an arbitrarily assumed Cartesian set of axes. Transition from one frame $\mathfrak{f}$ to another equally admissible one is accomplished by means of a one-to-one correspondence S in the domain of symbols x. One has to deal, therefore, with 4 kinds of objects: a set of symbols or cöordinates x, a group $\mathfrak{G}$ of transformations S of this set into itself, and further points P and frames $\mathfrak{f}$. Their connection is to be described thus: A point P relative to a frame $\mathfrak{f}$ determines a coördinate $x = (P, \mathfrak{f})$. Any two frames $\mathfrak{f}$, $\mathfrak{f}'$ determine a transformation S of our group $\mathfrak{G}$, such that the coördinate $x' = (P, \mathfrak{f}')$ of an arbitrary point P arises from its coördinate $x = (P, \mathfrak{f})$ by S, $x' = Sx$. With two given frames $\mathfrak{f}$, $\mathfrak{f}'$ the equation

$$(P, \mathfrak{f}) = (P', \mathfrak{f}') \tag{1}$$

defines an *automorphism* $P \rightleftarrows P'$ of the space. If the same S carries $\mathfrak{f}$, $\mathfrak{f}'$ into $\mathfrak{g}$, $\mathfrak{g}'$ respectively, one has, along with (1),

$$(P, \mathfrak{g}) = (P', \mathfrak{g}').$$

This shows that the automorphisms of our space form a group isomorphic with $\mathfrak{G}$; however, the isomorphic correspondence between the two groups depends on the frame of reference and is hence determined up to an arbitrary inner automorphism of the group. If in studying a given group $\mathfrak{G}$ of transformations $x' = Sx$ in a domain of symbols x one wishes to make use of a geometric nomenclature, it is quite fitting to *invent* a point space with its equally admissible frames to which the above scheme applies.

However, in one regard the scheme is still incomplete. Not only are the *points* of the space to be submitted to symbolical representation, but as has been emphasized by Plücker, other geometric entities also, e.g., the straight lines, may serve as spatial elements. Nay, in physics all sorts of physical quantities, velocities, forces, field strengths, electronic spins, etc., should be fixed by numerical symbols relative to a frame of reference. The law according to which the transformation S depends on the transition $\mathfrak{f} \to \mathfrak{f}'$ will then be determined by the type of the quantity in question, and will differ for points, lines, velocities, spins, etc. Only the elements s of the *abstract* group are tied up with the transitions in a manner independent of the type of quantity under consideration. After this correction, Klein's axiomatics looks as follows. (In its description I use the

language of physicists: instead of several points, I speak of *a* point which may assume several positions, or rather of a quantity, e.g., the electromagnetic field strength, capable of several values.)

A. *The "symbolic" part* (dealing with group elements and coördinates).

(1) Let there be given a set γ of elements called *group elements*. Each pair s, t of group elements shall give rise to a composite element ts. There shall be a unit element e satisfying $es = se = s$ and an inverse s^{-1} for each group element s: $s^{-1}s = ss^{-1} = e$. (The associative law is not explicitly required.)

(2) Let there be given a set of elements called coördinates x and a realization $\mathfrak{A}$: $s \to S$ of the group γ by means of one-to-one correspondences $x \to x' = Sx$ within that set.

B. *The "geometric" part* (dealing with frames and quantities).

(1) Any two *frames* $\mathfrak{f}$, $\mathfrak{f}'$ determine a group element s, called the *transition* from $\mathfrak{f}$ to $\mathfrak{f}'$. Vice versa, a group element s "carries" a frame $\mathfrak{f}$ into a uniquely determined frame $\mathfrak{f}' = s\mathfrak{f}$ such that the transition $(\mathfrak{f} \to \mathfrak{f}') = s$. The transition $\mathfrak{f} \to \mathfrak{f}$ is the unit element e, the transition $\mathfrak{f}' \to \mathfrak{f}$ the inverse element. If s, t are the transitions $\mathfrak{f} \to \mathfrak{f}'$, $\mathfrak{f}' \to \mathfrak{f}''$ respectively, then the composite ts is the transition $\mathfrak{f} \to \mathfrak{f}''$.

(2) A quantity q of the type $\mathfrak{A}$ is capable of different values. Relative to an arbitrarily fixed frame $\mathfrak{f}$ each value of q determines a coördinate x such that $q \to x$ is a one-to-one mapping of the possible values of q on the set of coördinates. The coördinate x' corresponding to the same arbitrary value q in any other frame $\mathfrak{f}'$ is linked to x by the transformation $x' = Sx$ associated with the transition $(\mathfrak{f} \to \mathfrak{f}') = s$ by the given realization $\mathfrak{A}$.

(1) refers to the space, (2) to a special quantity therein.

This sounds fairly general and abstract. As algebraists we are interested almost exclusively in the case where the realization of the group is a *representation* $s \to A(s)$ by linear transformations $A(s)$ in an n-dimensional vector space and where the coördinate is therefore a *k-vector*, i.e., any n-tuple of numbers $(x_1, \cdots, x_n)$. By "number" we mean here a number in an arbitrarily given field k. With this limitation we repeat once more our definition of a quantity:

A quantity q of type $\mathfrak{A}$ is characterized by a representation of γ in k, $s \to A(s)$, of a certain degree n. Each value of q relative to a frame $\mathfrak{f}$ determines a k-vector $(x_1, \cdots, x_n)$ such that under the transition s to another frame $\mathfrak{f}'$ the components x_i of q transform according to $A(s)$. [The representation $s \to 1$ of degree 1 is called the identical representation. A quantity of this type is a *scalar*.]

For the purposes of differential geometry this set-up is also of basic importance, though its does not tell the whole story. Here the procedure consists in associating with each point of the non-homogeneous "differential" manifold M a homogeneous Klein space of fixed type $\mathfrak{G}$ and in establishing transitions between these Klein spaces by moving around in M. For example, in a recent review of E. Cartan's method of répères mobiles in the Bulletin of the American Mathematical Society, I was able to show the adequacy of the axiomatic foundation as given here for his treatment of manifolds M, that are embedded in a Klein space,

by means of differential invariants. But I shall not enter into this subject here, my sole concern at present being algebraic invariants.

I denote by $\mathrm{P} = \mathrm{P}_n$ the "space" of n-dimensional k-vectors $(x_1, \cdots, x_n)$. A change of the vector basis in P transmutes $\mathfrak{A}$ into an equivalent representation $\mathfrak{A}'$. $\mathfrak{A}$ or the corresponding type of quantities is *reducible*, provided P has a linear subspace P′ invariant under all transformations $A(s)$ of the group $\mathfrak{A}$ which is neither the total P nor contains only the vector 0. By appropriate choice of the vector basis one then may split off a part of the components x_i such that these transform only among themselves. *Decomposition* occurs if P can be decomposed into two complementary invariant subspaces $\mathrm{P}_1 + \mathrm{P}_2$. This means that, relative to a suitable vector basis, the components break up into two classes

$$x_1, \cdots, x_l \mid y_1, \cdots, y_m \qquad (l + m = n), \tag{2}$$

the members of each transforming among themselves. The corresponding quantity consists of the juxtaposition of two quantities x and y which vary independently of each other. Thus one may look upon the electromagnetic four-potential together with the field strength as *one* quantity of $4 + 6 = 10$ components; but everybody will agree that this is a very artificial union. Looking from the other direction one will try and wish to *decompose every quantity into independent irreducible* ("primitive") *constituents*. For most groups, indeed for all which will engage our attention here, this is in fact possible. But the demonstration by algebraic means of the theorem of full reducibility is one of the chief goals of the theory.

Juxtaposition was defined thus: If the variables $x_1, \cdots, x_l$ are subject to the substitution A, and $y_1, \cdots, y_m$ to the substitution B, then the row (2) undergoes the substitution $A \dotplus B$. Another process of great importance is $\times$-multiplication: under the conditions just described, the lm products $x_i y_k$ undergo the substitution $A \times B$ which one calls the Kronecker product. Hence one may add and multiply representations $\mathfrak{A}$: $s \to A(s)$ and $\mathfrak{B}$: $s \to B(s)$ of the same group:

$$\mathfrak{A} + \mathfrak{B}: s \to A(s) \dotplus B(s); \qquad \mathfrak{A} \times \mathfrak{B}: s \to A(s) \times B(s);$$

or what is the same, one may add and $\times$-multiply quantities. In performing the second process, the representation spaces P and P′ of l and m dimensions over which the vectors x and y range, give rise to an lm-dimensional space PP′ which contains the vector $z = x \times y$ with the components

$$z_{ik} = x_i y_k .$$

In studying linear forms in PP′ one often finds it convenient to replace the most general vector z with lm independent components z_{ik} by the vector $x \times y$ with x_i and y_k as independent variables. This procedure is called the *symbolic method* in the theory of invariants. One of the most important problems for quantities is to decompose the product of two primitive quantities (or of two irreducible representations) into its irreducible constituents. Special cases will soon occupy us.

After all these preliminaries I shall finally say what an *invariant* is. I begin with the notion of a *vector invariant* which presupposes that we are given a group Γ of linear transformations A in an n-dimensional vector space P. Suppose we are given a form $f(x, y, \cdots)$, i.e., a homogeneous polynomial of certain degrees $\mu, \nu, \cdots$ in the components of each argument vector $x, y, \cdots$ which vary in P. The cogredient transformations

$$x' = Ax, \qquad y' = Ay, \qquad \cdots$$

change f into a new form $f' = Af$ defined by

$$f'(x', y', \cdots) = f(x, y, \cdots).$$

f is an (absolute) vector invariant with respect to the group Γ if $f = Af$ for all transformations A in Γ. A simple generalization of this elementary concept will introduce contravariant argument vectors $\xi, \eta, \cdots$ which undergo the transformation contragredient to A while the covariant arguments $x, y, \cdots$ are transformed by A.

To this elementary notion I oppose the *general notion of invariants* resting upon a given *abstract* group $\gamma = \{s\}$ and a series of given representations of γ,

$$\mathfrak{A}\colon s \to A(s), \qquad \mathfrak{B}\colon s \to B(s), \qquad \cdots$$

of degrees $m, n, \cdots$, respectively. A function $\varphi(\mathfrak{x}, \mathfrak{y}, \cdots)$ depending on an arbitrary quantity $\mathfrak{x}$ of type $\mathfrak{A}$, another quantity $\mathfrak{y}$ of type $\mathfrak{B}, \cdots$ will be expressed by a certain function F of the numerical vectors

$$x = (x_1, \cdots, x_m), \qquad y = (y_1, \cdots, y_n), \qquad \cdots$$

in terms of a given frame of reference $\mathfrak{f}$, and will be expressed by a certain function $F' = sF$ in terms of another frame $\mathfrak{f}'$ into which $\mathfrak{f}$ changes by the group element s. If $F' = F$ for all s, then φ is an invariant. If we make use of the numerical vectors and the given representations only, invariance may be simply stated by the equation

$$F(A(s)x, B(s)y, \cdots) = F(x, y, \cdots),$$

holding for all s in γ. Again, we limit ourselves to the case where F is a polynomial homogeneous in the components of each vector. Another way of expressing the same thing would be to say that *an invariant is a scalar depending on variable quantities of given types* $\mathfrak{A}, \mathfrak{B}, \cdots$.

One speaks of a *relative invariant* if $F' = \lambda \cdot F$ where the multiplier $\lambda = \lambda(s)$ depends on s only. $s \to \lambda(s)$ is then necessarily a representation of degree 1. More generally, a *covariant* is a quantity of a certain type $\mathfrak{H}\colon s \to H(s)$, depending on variable quantities $x, y, \cdots$ of given types $\mathfrak{A}, \mathfrak{B}, \cdots$.

After having fixed the concepts, we can now turn to the fundamental theorems concerning invariants. The *first main theorem* maintains that the invariants for a given group γ and a given set of its representations $\mathfrak{A}, \mathfrak{B}, \cdots$ have a *finite integrity basis*; i.e., one can pick out a finite number among them in terms of

which all these invariants are expressible in an integral rational manner. We do not know whether the proposition holds good for any group γ and representations $\mathfrak{A}$, $\mathfrak{B}$, $\cdots$. One has been able to prove it, however, in the most important cases, in particular for finite groups γ. After one has ascertained a finite complete set of basic invariants $J_1(x, y, \cdots), \cdots, J_h(x, y, \cdots)$, the second task is to survey all existing algebraic *relations* among them. A relation is a polynomial $R(t_1, \cdots, t_h)$ of h variables $t_1, \cdots, t_h$ which is turned identically into zero by the substitution

$$t_1 = J_1(x, y, \cdots), \cdots, t_h = J_h(x, y, \cdots).$$

The *second main theorem* states that one can find a finite number of relations of which all relations are algebraic consequences. This is merely a special case of Hilbert's universal proposition about the finiteness of an ideal basis for any ideal of polynomials. Indeed the relations form an ideal in the ring of all polynomials of $t_1, \cdots, t_h$. Thus the second main theorem is settled once for all and we shall pay little further attention to it. To be sure, in each single case the problem remains actually to ascertain an ideal basis of the relations.

I give two examples from the elementary domain of vector invariants. Let us deal with invariant forms depending on an arbitrary number of vectors $x^{(1)}, x^{(2)}, \cdots$ in the same space; invariance refers to a given group Γ of linear transformations in that space. If Γ is the group of all unimodular transformations, one gets an integrity basis by forming from the given argument vectors in all possible combinations the determinants $[xy \cdots z]$ of the components of n vectors $x, y, \cdots, z$. If contravariant arguments $\xi^{(1)}, \xi^{(2)}, \cdots$ are admitted, one must add the following two types

$$(\xi x) = \xi_1 x_1 + \cdots + \xi_n x_n$$

and $[\xi\,\eta \cdots \zeta]$. On the other hand, if Γ is the group of all orthogonal transformations, then the scalar products

$$(xy) \qquad \begin{Bmatrix} x = x^{(1)}, x^{(2)}, \cdots \\ y = x^{(1)}, x^{(2)}, \cdots \end{Bmatrix}$$

of the argument vectors constitute an integrity basis for invariants. Surprisingly enough the last result holds good even when the underlying number field is any field of characteristic zero in the sense of abstract algebra. Since the construction of suitable Cartesian coördinate systems to which the proofs resort depends on laying off a given segment on a given line, one would have expected the result to be restricted to "*Pythagorean*" *fields*. As one knows, a field is called *real* (Artin-Schreier) provided a square sum never vanishes unless each term vanishes. I name a real field Pythagorean if the square sum of two numbers is always a square. All relations between scalar products are in the case of the orthogonal group consequences of the relations of the following type:

$$\begin{vmatrix} (xx) & (xx') & \cdots & (xx^{(n)}) \\ \cdots & \cdots & \cdots & \cdots \\ (x^{(n)}x) & (x^{(n)}x') & \cdots & (x^{(n)}x^{(n)}) \end{vmatrix} = 0.$$

(Second main theorem for orthogonal vector invariants.)

By means of his theorem about polynomial ideals Hilbert had reduced the general proof of the first main theorem to the construction of a linear operator ω working on polynomials $F(x, y, \cdots)$ and having the following two properties:

$$\omega(1) = 1, \qquad \omega(F \cdot J) = \omega(F) \cdot J \tag{3}$$

whenever J is an invariant. If γ is a compact Lie group, one can follow a procedure inaugurated by Adolf Hurwitz and define an *invariant measure of volumes* on γ by means of which one is able to form the average $\mathfrak{M}_s\{\psi(s)\}$ of any continuous function $\psi(s)$ on γ. One sees at once that

$$\omega(F) = \mathfrak{M}_s(sF)$$

is a process of the desired nature. By this topological method which necessarily presupposes the continuum K of all real numbers as reference field, one succeeds in proving the first fundamental theorem for any compact Lie group. I mention the instance of the real orthogonal group in K. By the same method I. Schur succeeded in carrying over from finite to compact Lie groups Frobenius' theory of group representations, in particular, the orthogonality relations for the representing matrices and their characters, while the speaker, together with F. Peter, established the completeness relation.

A. Haar freed the definition of the volume measure of the awkward differentiability conditions imposed by the Lie nature of the group. H. Bohr's theory of almost periodic functions could be interpreted as the simplest example of a similar theory for open, non-compact groups, namely, for the group of translations of a straight line. With the theory of compact groups and Bohr's example of a non-compact group before his eyes, von Neumann established the theory of almost periodic representations, their orthogonality and completeness, for any group whatsoever. Hence the first main theorem for invariants is proved for each group as long as we restrict ourselves to quantities $x, y, \cdots$ as arguments whose types are described by almost periodic representations.

All this sounds as if we could rest as God did after the sixth day of creation, finding that it was very good! But now enters the snake into the paradise. Let us once more envisage the classical case of the group L^r of all real unimodular transformations A in n dimensions. Not one of the representations with which the classic theory of invariants deals, not even the representation $A \rightarrow A$, is almost periodic! Thus the "almost periodic" theory fails just in the most important and natural cases. Nevertheless it has been possible to make the theory of compact groups fruitful for all semi-simple Lie groups by what I have called the *unitarian trick*. For the group L^r it consists in first extending L^r to embrace all unimodular transformations with *complex* coefficients and then limiting oneself within this wider group L^c to the *unitary* operations. By following Lie's fundamental suggestion and going back to the infinitesimal elements of a group, one linearizes and thereby algebraizes all problems concerning structure, representations and invariants of a group; and then such reality restrictions as the two encountered above, either to real coefficients or to the

unitary subgroup, become irrelevant. Hence each of these subgroups can stand for the other, and one of them, namely, the unitary subgroup, is compact and thus accessible to the integration method. In the linkage between the infinitesimal and the total group a topological element is involved; but I shall not dwell here on this subtle point. Anyway I have been able to show that the unitarian trick is effective with all semi-simple Lie groups, and thus not only to confirm by a combination of the infinitesimal and integral methods the results derived in a purely infinitesimal manner by E. Cartan for the irreducible representations of the semi-simple groups, but also to supplement them by the theorem of full reducibility and explicit formulas for their characters. At the same time *the first main theorem for invariants was thus secured for all semi-simple groups.*

The problem naturally puts itself: to corroborate by direct and explicit algebraic construction these results first obtained in a transcendental way. If one succeeds, one may hope at the same time to remove the bond by which the topological approach ties these results to the field K of real numbers and to extend them to any field in the abstract algebraic sense, at least to any field of characteristic zero. This is a goal at which I have aimed for many years, though not at all with the necessary persistence and singleness of purpose. So many other mathematical things have diverted my interests, and the whirlwind of political events has had a most disturbing effect on my concentration. However, younger men came to my aid, above all Richard Brauer, to whom I owe the most essential link in the chain of the algebraic theory. At present I have come to a certain end, or at least to a certain halting point, from which it seemed profitable to look back upon the track so far pursued, and this is what I tried to do in my recent book *The Classical Groups, their Invariants and Representations.* The most important simple groups in the field of all complex numbers are: the group $L(n)$ of all (non-singular or merely of all unimodular) linear transformations in n dimensions, the group $O(n)$ of all (or all proper) orthogonal transformations in n dimensions, and the group $Sp(n)$ of all linear transformations in $n = 2\nu$ dimensions leaving invariant a non-degenerate skew-symmetric bilinear form. The last I have christened the symplectic group. These are even the only ones, apart from 5 quite singular exceptional groups. I shall deal exclusively with these groups $L(n)$, $O(n)$, $Sp(n)$. For their investigation a finite group, the group of all permutations, must be drawn in, and one could also include the alternating group of permutations. These groups are in my mind when I speak of *classical groups.* We are first engaged in algebraically constructing the *possible types of quantities* under their reign.

Again we start with the universal linear group $L(n)$, an arbitrary element of which we denote by A:

$$x_i' = \sum a(ik)x_k \qquad (i, k = 1, \cdots, n).$$

You all know what a tensor of rank r is. It has n^r components $\mathrm{t}(i_1 i_2 \cdots i_r)$ labeled by r indices $i_1, i_2, \cdots, i_r$ ranging from 1 to n; under the influence of

the transformation A of the coördinates in the underlying vector space these components are transformed according to the substitution

$$\Pi_r(A) = A \times A \times \cdots \times A \qquad (r \text{ factors})$$

or more explicitly

$$(4) \qquad t'(i_1 \cdots i_r) = \sum_{k_1, \cdots, k_r} a(i_1k_1) \cdots a(i_rk_r) \cdot t(k_1 \cdots k_r).$$

The generic tensor of rank r is the quantity arising by r-fold $\times$-multiplication of the quantity vector. But the space P^r of all tensors is not irreducible under the group $\Pi_r(L)$ consisting of the substitutions $\Pi_r(A)$ which are induced in tensor space by the elements A of $L(n)$, whereas the words symmetric tensor, skew-symmetric tensor, indicate irreducible quantities. The tensor space P^r must therefore be split into irreducible invariant parts by imposing symmetry conditions upon the tensors. The possibility of doing so is based on the fact that one can perform an arbitrary permutation p on the r indices or arguments $i_1, \cdots, i_r$, whereby t changes into another tensor pt. In this way enters the group π_r of permutations p of r figures $1, \cdots, r$. Associating the transition $t \to pt$ with p defines a representation of π_r by linear transformations in P^r. But why is it that these permutation operators are of importance for the decomposition of tensor space into invariant subspaces? One understands this if one replaces the group $\Pi_r(L)$ of the substitutions (4) to which the tensor space is submitted by its *enveloping algebra*, containing all those substitutions which can be gained by linearly combining any finite number of substitutions of the group $\Pi_r(L)$. It is easily seen that the enveloping algebra consists of all linear substitutions $t \to Ht$ commuting with the permutations $p: t \to pt$. The group π_r of permutations may also conveniently be replaced by the enveloping algebra, i.e., by the corresponding group ring whose elements

$$\sum_p \alpha(p) \cdot p \qquad [\alpha(p) \text{ numbers}]$$

may be interpreted as "symmetry operators" working on tensors.

The general situation under which our problem is naturally to be subsumed is now this: Instead of the tensor space we consider an arbitrary vector space P whose vectors are called t; there is given a finite group $\gamma = \{p\}$ and a representation of γ in P representing the abstract group element p by a linear substitution $t \to pt$. We are interested in the algebra $\mathfrak{A}$ of linear operators $t \to Ht$ commuting with all operators $t \to pt$ of γ. The regular representation r of a finite group γ or of its group ring $(\gamma) = \rho$ has ρ itself as its representation space, representing any element a of ρ by the transformation $x \to ax$ of ρ into itself. By a well-known theorem due to Maschke the regular representation of γ is fully reducible; this holds good in any field, unless it is of a prime characteristic dividing the order of γ. We take into account only fields of characteristic zero. A thoroughly elementary method permits establishment of a complete parallelism between the subspaces of ρ invariant under r on the one side and the subspaces

of P invariant under the algebra $\mathfrak{A}$ on the other side. The parallelism is faithful with respect to addition and the relation of being contained for invariant subspaces, and also with respect to equivalence under their respective operator algebras $\mathfrak{r}$ and $\mathfrak{A}$.

For the symmetric group π_r one knows how to carry out the decomposition into irreducible invariant subspaces by means of the symmetry operators which were invented by A. Young and later, under the leadership of E. Wigner, have found such surprising applications in quantum mechanics. Let us attach the word *quantics*, originally coined by Cayley, to the quantities which one prepares in this way from the material of tensors under the rule of the full linear group $L(n)$. The domain of quantics is closed with respect to the two most important operations: (1) $\times$-multiplication of two quantics followed by decomposition into irreducible constituents, (2) transition from a representation to its contragredient. Each Young operator and hence each quantic[1] is characterized by a partition of the rank number r into n integral summands

$$r = r_1 + r_2 + \cdots + r_n \quad (r_1 \geqq r_2 \geqq \cdots \geqq r_n \geqq 0).$$

We represent this partition by a symmetry diagram whose rows have the lengths $r_1, r_2, \cdots, r_n$. Example:

$r_1 = 7$ ○○○○○○○
$r_2 = 5$ ○○○○○
$r_3 = 5$ ○○○○○
$r_4 = 2$ ○○
$r_5 = 1$ ○

If one wishes to employ a similar method for the orthogonal and the symplectic groups one has first to get hold by a simple description of the enveloping algebra of the substitutions $\Pi_r(A)$ induced in tensor space by the elements A of these more limited groups $O(n)$ and $Sp(n)$. The problem is not as trivial by far as in the former case of the full linear group $L(n)$, and R. Brauer succeeded in solving it only by resorting to the general theory of matric algebras. If one is given an algebra $\mathfrak{A}$ of linear substitutions or matrices A in a certain vector space P, then the matrices B commuting with all matrices A of the set $\mathfrak{A}$ form in their turn an algebra $\mathfrak{B}$ which I call *the commutator algebra* of $\mathfrak{A}$. The key principle asserts that if $\mathfrak{A}$ is fully reducible, the commutator algebra of the commutator algebra of $\mathfrak{A}$ is not larger than $\mathfrak{A}$ as one might expect, but coincides with $\mathfrak{A}$. This principle holds in any field. It is the crowning result of a theory of matric algebras based on this fundamental advice due to I. Schur: along with a given matric algebra, always consider its commutator algebra. Unable to refer to any other place, I had to incorporate in my book this theory which has become a central issue in the whole non-commutative algebra. Perhaps many a reader will find such a concrete treatment in terms of matrices more easily accessible than the abstract handling of semi-simple rings.

[1] We disregard here a slight modification necessary to work the process (2).

If one replaces the group $\Pi_r(O)$ induced by the orthogonal group $O(n)$ in tensor space by its enveloping algebra thus determined, one succeeds again in decomposing the most general tensor into quantics with respect to the orthogonal group. The primitive quantics to which one finds oneself reduced differ from those for the full linear group in two essential regards: (1) The Young symmetry operator is applied not to an arbitrary tensor, but to the most general tensor whose $\frac{1}{2}r(r-1)$ traces vanish; the (12)-trace of a tensor $\mathfrak{t}(i_1 i_2 i_3 \cdots i_r)$ being given by

$$\mathfrak{t}_{12}(i_3 \cdots i_r) = \sum_i \mathfrak{t}(i\, i\, i_3 \cdots i_r)$$

(process of *Verjüngung*). (2) While all symmetry diagrams whose first column had a length $\leqq n$ were admitted for the full linear group, only those occur here whose first two columns have a total length $\leqq n$. Similar results obtain for the symplectic group which in many regards is easier to handle than the orthogonal group.

The algebraic concept of invariants which we adopt for the classical groups is that of a scalar depending on one or more independent variable *quantics*. Arbitrary *forms* such as occurred as arguments in the classical theory of invariants are identical with arbitrary symmetric tensors, and under the reign of $L(n)$ these are quantics which correspond to a symmetry diagram of one row (the length of the row being the order of the form). Let us first stick again to $L(n)$! In all textbooks on our subject one is told a lot about the so-called *symbolic method* which reduces *form invariants* to *vector invariants*. On the basis of the above analysis one shows readily that the method still works for invariants of variable quantics of any type. However, since the number of argument vectors to be introduced depends on the degree of the invariant under consideration with respect to the variable quantics, the reduction to vector invariants is by no means sufficient without further resources for a proof of the first main theorem. Certainly the importance of the symbolic method whose formal elegance nobody will deny has been greatly exaggerated. I consider it one of the more glaring examples of the power of tradition and inertia in mathematics that the elementary textbooks on invariants up to this day deal almost exclusively with this method and its applications. Frequently quite different approaches, e.g., irrational methods, lead much faster to the goal. Hilbert's general proof dealing with form invariants of the group $L(n)$ is based on his general proposition about a finite ideal basis for polynomial ideals and employs as the above mentioned process ω with the properties (3) Cayley's purely algebraic Ω-process. The method goes through in any field of characteristic zero even if quantics instead of forms are the arguments.

The projective geometers were able to cope with affine and metric geometry by adjoining some entities given once for all which they called the *absolute*: the plane at infinity and the absolute involution. In the same line of ideas relativity theory has found it convenient to treat the metric vector space as an affine vector space in which a positive quadratic form is appointed as metric

ground form. Is this standpoint justified as far as invariants are concerned? Let variable forms $u, v, \cdots$ be our arguments in the invariants $J(u, v, \cdots)$ which we envisage. Then the question means whether any *orthogonal* invariant $J(u, v, \cdots)$ arises by the Cartesian specialization $g_{ik} = \delta_{ik}$ from a L-invariant $J(g_{ik}; u, v, \cdots)$ involving besides the arguments $u, v, \cdots$ a variable covariant quadratic form

$$\sum_{i,k} g_{ik}\xi_i\xi_k. \tag{5}$$

That it can answer this question in the affirmative goes to the credit of the symbolic method and is its highest triumph. Indeed, each orthogonal vector invariant is expressible in terms of the scalar products (xy) of the vectorial arguments, and if we replace the scalar product

$$x_1y_1 + \cdots + x_ny_n \tag{6}$$

by the L-invariant formation

$$-\begin{vmatrix} g_{11} & \cdots & g_{1n} & x_1 \\ \cdots & \cdots & \cdots & \cdots \\ g_{n1} & \cdots & g_{nn} & x_n \\ y_1 & \cdots & y_n & 0 \end{vmatrix}$$

which depends on (5) besides the two vectors x and y and changes into (6) by the specialization $g_{ik} = \delta_{ik}$, then we attain our ends, first for vector invariants and then, owing to the symbolic treatment, for form invariants. The procedure remains applicable even for arbitrary quantics, although the quantics for the group L split into more primitive quantics under the orthogonal subgroup O. In this purely algebraic way based on the adjunction argument we master the orthogonal and the symplectic invariants. This procedure has even stood the test in certain special cases where the statement of full reducibility breaks down.

In these days the angel of topology and the devil of abstract algebra fight for the soul of each individual mathematical domain. As to the decompositions into invariant subspaces whose algebraic construction has here been indicated, one would like to know in some explicit way with which *multiplicity* each of the inequivalent irreducible constituents occurs. This question is answered most readily if one replaces the representations by their *characters*. Explicit formulas for the characters are much more easily obtained by the *integral-topological approach*. The identification of the characters thus derived with the algebraically constructed representations to which they correspond causes a little headache; however, if this has been remedied one has seized upon a result which by its very nature is independent of the nature of the reference field, and that in spite of the fact that the topological method operates in the field K of all real numbers. Thus we are face to face with a peculiar application of analysis to purely algebraic problems whose stage is set in an arbitrary field. The multi-

plicities just mentioned yield at the same time formulas for the explicit enumeration of invariants and covariants. This field has recently been tilled with high success by Professor Murnaghan.

I must forego giving examples of such enumerating formulas. Instead I prefer to mention a by-product of the algebraic investigation. In the n^2-dimensional space of all matrices $A = \| a_{ik} \|$ the equations

$$(7) \qquad \sum_k a_{ik} a_{jk} - \delta_{ij} = 0, \qquad\qquad \det(a_{ik}) - 1 = 0$$

define a certain algebraic manifold, the proper orthogonal group O^+. *This manifold is irreducible*, or, what is the same, the ideal of all polynomials $\Phi(a_{ik})$ vanishing on $O^+(n)$ is a prime ideal. The polynomials constituting the left members of the equations (7) form an ideal basis for our ideal, and Cayley's rational parametrization of orthogonal substitutions A,

$$A = (E - S)(E + S)^{-1},$$

in terms of an arbitrary skew-symmetric matrix S yields a generic zero (allgemeine Nullstelle) of the ideal, provided the elements s_{ik} $(i < k)$ of the skew matrix $S = \| s_{ik} \|$ are treated as indeterminates. All this holds good in any field of characteristic zero.

I hope my sketch has shown how closely the investigation of the invariants of a group is tied up with the ascertainment of its representations. This connection with the general theory of representations and of matric algebras has carried new life-blood into the older theory of invariants which thus has joined the modern forward movement of algebra and now participates in its general conceptual structure. I feel bound to add a personal confession. In my youth I was almost exclusively active in the field of analysis; the differential equations and expansions of mathematical physics were the mathematical things with which I was on the most intimate footing. I have never succeeded in completely assimilating the abstract algebraic way of reasoning, and constantly feel the necessity of translating each step into a more concrete analytic form. But for that reason I am perhaps fitter to act as intermediary between old and new than the younger generation which is swayed by the abstract axiomatic approach, both in topology and algebra.

In closing I should like to point out a few lines of probable *further advance*. First, one naturally wishes to do all things also in a field of prime characteristic. Secondly, it is desirable to find *all* inequivalent irreducible representations in an arbitrarily given field; it is doubtful enough that they are exhausted by the quantics, though these form a class of quantities algebraically closed in a certain sense. If one replaces a continuous group by its infinitesimal elements, one has to deal with a *Lie algebra* and one will ask for its representations and invariants. The classical groups together with the 5 exceptional groups mentioned above yield all the simple Lie-algebras in the field of complex numbers, or in any

algebraically closed field. However, this does not remain true in an arbitrary field. In this question Landherr, A. A. Albert, Jacobson and Zassenhaus have recently made much headway. So I am confident that in a few years a younger algebraist will be able to write a similar book dealing comprehensively with the representations and invariants of all semi-simple Lie algebras in an arbitrary field.

INSTITUTE FOR ADVANCED STUDY.

ADDITIVITY AND CONTINUITY OF PERSPECTIVITY

By Israel Halperin

Introduction. If L is the set of linear subspaces of a projective geometry and $D(a)$ is the common dimension function[1] defined for the elements of L, it is easily shown that

$$D(a + b) + D(ab) = D(a) + D(b) \qquad \text{for all } a, b \text{ in } L. \tag{1}$$

But if $D(a)$ is defined[2] to be the common dimension of a plus 1, then $D(a)$ satisfies not only (1) but also[3] $D(0) = 0$ and hence

$$D(a + b) = D(a) + D(b) \qquad \text{if } ab = 0. \tag{2}$$

More generally, it will follow that for every finite m

$$D\left(\sum_{r=1}^{m} a_r\right) = \sum_{r=1}^{m} D(a_r) \tag{3}$$

if $a_1, \cdots, a_m$ are linearly independent. Since in a given projective geometry there is a finite upper bound for the number of elements which can all be different from 0 and be linearly independent, there is no point in considering (3) for non-finite sums.

J. von Neumann has given a remarkable set of axioms defining a class of geometries which includes all projective geometries as well as a new type of geometry which he has named continuous geometry.[4] He has shown, too, that in each of these geometries there exist a dimension function $D(a)$ and a concept of independence such that (3) holds for all sets of independent elements whether their number is finite or not. And in continuous geometries there do exist countable sets of elements which are all different from 0 and are independent, but there are no such non-countable sets. In all of these geometries two elements a, b have the same dimension if and only if a can be mapped into b by a perspectivity (denoted by the symbols $a \sim b$).[5]

Among the axioms given by von Neumann to define projective and con-

Received September 28, 1938; in revised form, March 23, 1939.

[1] That is, dimension of a point = 0, dimension of a line = 1, etc. The subspace of L which does not contain any points has common dimension = -1.

[2] It would be even more suggestive, from the point of continuous geometries, to define $D(a)$ to be [(common dimension of a) + 1]/[(common dimension of L) + 1].

[3] The same symbol will be used to denote both the number 0 and the empty subset of L, but there should be no confusion.

[4] See [3], [4], [5]. (Numbers in brackets refer to the bibliography at the end of the paper.)

[5] The actual construction of the dimension function is closely related to this property.

tinuous geometries there is an axiom of irreducibility.[6] This axiom of irreducibility is essential for the construction of the dimension function, but it is not needed in the definition of perspectivity nor in the definition of independence. This makes it possible to formulate (3) in a way that has meaning in all reducible geometries, that is, in all systems which satisfy all of von Neumann's axioms with the possible exception of the irreducibility axiom. The new formulation states that for every set I of indices

(4) If a_α, $\alpha \in I$, are independent and b_α, $\alpha \in I$, are independent, and if $a_\alpha \sim b_\alpha$ for every $\alpha \in I$, then $\sum_{\alpha \in I} a_\alpha \sim \sum_{\alpha \in I} b_\alpha$;

that is, that perspectivity is unrestrictedly additive in reducible geometries. (4) is significant for all sets I of indices since in reducible geometries there can exist sets of elements of any given cardinal power which are all different from 0 and are independent.

The present paper will establish this unrestricted additivity of perspectivity as well as the continuity of perspectivity (described below). These properties of perspectivity have been established previously by the writer for countable sets of elements,[7] and the methods used will be simplified here considerably and extended to give the general theorems. In §1 the axioms and definitions to be used will be stated precisely. In §2 the theory of independence and other preliminary theory of von Neumann will be recalled. The continuity and additivity of perspectivity will be established, as stated, in §3 and the axioms on which the proofs depend will be somewhat weakened in §4.

1. Axioms and definitions.[8]

Definition 1.1. A reducible geometry is a system L of elements $a, b, c, \cdots$ with a relation $a \leqq b$ (equivalently, $b \geqq a$) defined for certain pairs of elements of L and satisfying the following axioms.

Axiom I. Partial ordering.

I_1. *$a \leqq b$, $b \leqq c$ together imply $a \leqq c$, and*

I_2. *$a \leqq b$, $b \leqq a$ are together equivalent to $a = b$.*

Axiom II. Completeness. *For every subset of elements a_α, $\alpha \in I$, there exists*

II_1. *a sum element $\sum_{\alpha \in I} a_\alpha$ such that for every a in L, $a \geqq \sum_{\alpha \in I} a_\alpha$ if and only if $a \geqq a_\alpha$ for every $\alpha \in I$.*

II_2. *an intersection element $\prod_{\alpha \in I} a_\alpha$ such that for every a in L, $a \leqq \prod_{\alpha \in I} a_\alpha$ if and only if $a \leqq a_\alpha$ for every $\alpha \in I$.*

[6] See [3], Axiom VI.

[7] See [1], §6.

[8] The axioms and definitions of this section will be found in slightly different but equivalent form in [3].

AXIOM III. CONTINUITY. *For every ordinal* Ω *and for every set of elements* a_α, $\alpha < \Omega$, *and for every* b,

III$_1$. $b \sum_{\alpha<\Omega} a_\alpha = \sum_{\gamma<\Omega} (b \sum_{\alpha\leqq\gamma} a_\alpha)$.

III$_2$. $b + \prod_{\alpha<\Omega} a_\alpha = \prod_{\gamma<\Omega} (b + \prod_{\alpha\leqq\gamma} a_\alpha)$.

AXIOM IV. MODULARITY. *For all* a, b, c *in* L, $(a + b)c = \{a + (a + c)b\}c$, *or what is equivalent,* $a \leqq c$ *implies* $(a + b)c = a + bc$ *for all* b.

AXIOM V. COMPLEMENTATION. *For any* a, b, c *in* L *such that* $a \leqq b \leqq c$, *there exists at least one element, which may be denoted by* $[c - b]_a$, *such that* $b + [c - b]_a = c$ *and* $b[c - b]_a = a$.

Remark 1. This set of axioms is self-dual, that is, it is unchanged if $\leqq$, $\sum$, $\prod$ are everywhere replaced by $\geqq$, $\prod$, $\sum$, respectively. Hence any theorem which holds for $\leqq$, $\sum$, $\prod$ will hold also for the dual theorem which arises from such replacements.

Remark 2. If a_α ranges over all the elements of L, the elements $0 = \prod_\alpha a_\alpha$ and $1 = \sum_\alpha a_\alpha$ will have the property that, for every a in L, $0 \leqq a \leqq 1$. We shall frequently write $[c - b]$ in place of $[c - b]_0$.

DEFINITION 1.2. A set of elements a_α, $\alpha \in I$, is said to be independent, written $(a_\alpha, \alpha \in I)\perp$, if $a_\alpha(\sum_{\beta\in I,\beta\neq\alpha} a_\beta) = 0$ for every $\alpha \in I$. We shall frequently write $\sum_{\alpha\in I} \oplus a_\alpha$ in place of $\sum_{\alpha\in I} a_\alpha$ to denote that the a_α, $\alpha \in I$, are independent.

DEFINITION 1.3. Two elements a, b are said to be perspective, written $a \sim b$, if there exists any element c (called an axis) such that $a + c = b + c$ and $ac = bc$. If $a \sim b_1$ for some $b_1 \leqq b$, we write $a \propto b$.

DEFINITION 1.4. If a_α is defined for all $\alpha <$ some limit ordinal Ω, and $\prod_{\alpha<\Omega} \sum_{\beta\geqq\alpha} a_\beta = \sum_{\alpha<\Omega} \prod_{\beta\geqq\alpha} a_\beta = a$, we write $\lim_{\alpha\to\Omega} a_\alpha = a$.

Remark. If $a_\alpha \leqq a_\beta$ for all $\alpha \leqq \beta < \Omega$, $\lim_{\alpha\to\Omega} a_\alpha$ exists and is equal to $\sum_{\alpha<\Omega} a_\alpha$. If $a_\alpha \geqq a_\beta$ for all $\alpha \leqq \beta < \Omega$, $\lim_{\alpha\to\Omega} a_\alpha$ exists and $= \prod_{\alpha<\Omega} a_\alpha$.

2. Preliminary theory.[9]

THEOREM 2.1. *If* $(a_\alpha, \alpha \in I)\perp$ *and* $b_\alpha \leqq a_\alpha$ *for all* $\alpha \in I$, *then* $(b_\alpha, \alpha \in I)\perp$.

[9] The following theory of independence and perspectivity will be found in [4]. Our definition of $\sim$ is equivalent to the definition used in [4] by the Theorems 2.3 and 2.2 of [4], Part I. Our Theorems 2.2, 2.4, 2.6, 2.8, 2.10, 2.11, 2.12, 2.13 are identical with, or follow immediately from, the Theorems 2.3, 2.6, 2.4, 3.6, 3.3, 3.4, 3.8, 4.4, respectively, of [4], Part I. Our Theorem 2.7 is implied by the proof of Lemma 3.3 of [4], Part III, and is given explicitly as Lemma 3.3 in [2]. The transitivity of perspectivity, the first part of our Theorem 2.14, is the Theorem 2.3 of [4], Part III, and is proved with the use of only the "countable" axioms in [1]; the rest of our theorem follows as shown in Theorem 2.5 of

THEOREM 2.2. *$(a_\alpha, \alpha \in I)\perp$ if and only if $(a_\alpha, \alpha \in J)\perp$ for every finite subset J of I.*

THEOREM 2.3. *For every limit ordinal Ω, $(a_\alpha, \alpha < \Omega)\perp$ if and only if $(a_\beta, \beta < \alpha)\perp$ for all $\alpha < \Omega$.*

THEOREM 2.4. *If $(a_\alpha, \alpha \in I)\perp$, if J_β, $\beta \in P$, are a family of subsets of I and J is the set of indices common to all J_β, $\beta \in P$, then $\prod_{\beta \in P} (\sum_{\alpha \in J_\beta} a_\alpha) = \sum_{\alpha \in J} a_\alpha$.*

THEOREM 2.5. *If $(a_\alpha, \alpha \in I)\perp$ and J_β, $\beta \in P$, are mutually exclusive subsets of I, and if $b_\beta = \sum_{\alpha \in J_\beta} a_\alpha$, then $(b_\beta, \beta \in P)\perp$.*

THEOREM 2.6. *If $(a_{\alpha,\beta}, \beta \in I_\alpha)\perp$ for every $\alpha \in P$, if $a_\alpha = \sum_{\beta \in I_\alpha} a_{\alpha,\beta}$ and if $(a_\alpha, \alpha \in P)\perp$, then $(a_{\alpha,\beta}, \beta \in I_\alpha, \alpha \in P)\perp$.*

THEOREM 2.7. *If for some ordinal Ω decompositions $a_\alpha = a_{\alpha+1} \oplus a'_\alpha$ are defined for all $1 \leqq \alpha < \Omega$, and if $a_\gamma = \prod_{\alpha<\gamma} a_\alpha$ for all limit ordinals $\gamma \leqq \Omega$, then $a_1 = a_\Omega \oplus \sum_{\alpha<\Omega} \oplus a'_\alpha$.*

THEOREM 2.8. *If $(a_\alpha, b_\beta, \alpha < \Omega, \beta < \Omega)\perp$, and if $a_\alpha \sim b_\alpha$ for all $\alpha < \Omega$, then $\sum_{\alpha<\Omega} \oplus a_\alpha \sim \sum_{\alpha<\Omega} \oplus b_\alpha$.*

THEOREM 2.9. *If $a \sim b$ and $c \geqq a + b$, there exists an element x such that $a \oplus x = b \oplus x = c$.*

THEOREM 2.10. *If $a \sim b$ and $a_1 \leqq a$, then $a_1 \propto b$.*

THEOREM 2.11. *$a \sim b$, $b \sim c$ and $(a, b, c)\perp$ imply that $a \sim c$.*

THEOREM 2.12. *If $(a_n, n = 1, 2, \cdots)\perp$ and $a_n \sim a_{n+1}$ for all n, then $a_n = 0$ for all n.*

THEOREM 2.13. *$a \propto b$, $b \propto a$ are together equivalent to $a \sim b$.*

THEOREM 2.14. *If $a \sim b$ and $b \sim c$, then $a \sim c$. If $a \propto b$ and $b \propto c$, then $a \propto c$. The dual to $a \propto b$ is equivalent to $b \propto a$.*

3. Continuity and additivity of perspectivity.

THEOREM 3.1. *If $ab = 0$, then $(a + c)b \propto c$ for all c in L.*

Proof. Replacing a by $[a - ac]$, we may assume that $ac = 0$. Then $(a + c)b \sim (a + b)c$ with axis a since

$$a(a + c)b = ab = 0, \qquad a(a + b)c = ac = 0,$$

[4], Part III. Finally, our Theorems 2.1, 2.5 follow easily from the definition of $\perp$, our Theorem 2.3 follows easily from our Theorem 2.2, and our Theorem 2.9 follows from the Theorem 3.9 and the definition of $\sim$ of [4], Part I.

and, using the modular axiom,

$$a + (a + c)b = (a + b)(a + c) = a + (a + b)c.$$

Since $(a + b)c \leqq c$, we have $(a + c)b \propto c$ as required.

THEOREM 3.2.[10] *If $a \oplus b \propto c \oplus d$ and $c \propto a$, then $b \propto d$.*

Proof. Since perspectivity is transitive (Theorem 2.14), we may assume that $a + b \leqq c + d$. Replacing a by $[(c + d) - (a + b)] + a$, we may assume that $a \oplus b = c \oplus d$. If now $c \sim a_1 \leqq a$, we can replace a by a_1 and b by $b + [a - a_1]$. Thus we may assume that $a \sim c$, and hence by Theorem 2.9, $a \oplus x = c \oplus x = a \oplus b = c \oplus d$ for some x. Then $b \sim x$, $x \sim d$, and Theorem 2.14 implies that $b \sim d$. The theorem follows from this.

THEOREM 3.3.[10] *If $a \propto c$, $b \propto d$, and $ab = cd = 0$, then $a \oplus b \propto c \oplus d$. In particular, if $a \sim c$ and $b \sim d$, then $a \oplus b \sim c \oplus d$.*

Proof. Let $e = [(a + b + c + d) - (a + b)]$, $f = [(a + b + c + d) - (c + d)]$. Then $e \oplus a \oplus b = f \oplus c \oplus d$. Since $b \propto d$, Theorem 3.2 implies that $f \oplus c \propto e \oplus a$. Since $a \propto c$, Theorem 3.2 implies that $f \propto e$. From $e \oplus (a \oplus b) = f \oplus (c \oplus d)$ we finally derive by Theorem 3.2 that $a \oplus b \propto c \oplus d$, as required.

If $a \sim c$ and $b \sim d$, we can derive that $a \oplus b \propto c \oplus d$, and in the same way, $c \oplus d \propto a \oplus b$. From Theorem 2.13 it then follows that $a \oplus b \sim c \oplus d$.

THEOREM 3.4. *If $a \propto b + c$, then $a = a_1 \oplus a_2$ with $a_1 \propto b$ and $a_2 \propto c$.*

Proof. Since perspectivity is transitive, we may suppose that $a \leqq b + c$. Set $a_1 = ab$ and $a_2 = [a - a_1]$. Then $a = a_1 \oplus a_2$. Now $a_1 \leqq b$ implies $a_1 \propto b$, and $a_2 b = 0$, $a_2 \leqq b + c$ imply, by Theorem 3.1, $a_2 \propto c$. This proves the theorem.

THEOREM 3.5. *If $a \propto b_\alpha$ and $b_\alpha \geqq b_\beta$ for all $1 \leqq \alpha \leqq \beta < \Omega$, then $\prod_{\alpha<\Omega} b_\alpha = 0$ implies that $a = 0$.*

Proof. 1. If Ω has a predecessor λ, then $b_\lambda = \prod_{\alpha<\Omega} b_\alpha = 0$ and $a \propto b_\lambda$, and this implies that $a = 0$, as required. Hence we may assume that Ω is a limit ordinal. Replacing b_γ by $\prod_{\alpha<\gamma} b_\alpha$ for all limit ordinals $\gamma \leqq \Omega$, we may also suppose that $b_\gamma = \prod_{\alpha<\gamma} b_\alpha$ for all limit ordinals $\gamma \leqq \Omega$.

2. Since perspectivity is transitive, we may assume that $a \leqq b_1$. Now define $a_\alpha = ab_\alpha$, $a'_\alpha = [a_\alpha - a_{\alpha+1}]$. Then $a_\alpha = a_{\alpha+1} \oplus a'_\alpha$ for all $\alpha < \Omega$, and $a_\gamma = ab_\gamma = a(\prod_{\alpha<\gamma} b_\alpha) = \prod_{\alpha<\gamma} (ab_\alpha) = \prod_{\alpha<\gamma} a_\alpha$ for all limit ordinals $\gamma \leqq \Omega$. This implies by Theorem 2.7 (use the fact that $a_\Omega \leqq b_\Omega = 0$) that $a = a_1 = \sum_{\alpha<\Omega} \oplus a'_\alpha$.

[10] Theorems 3.2 and 3.3 are the Theorem 2.4 of [4], Part III, slightly strengthened.

3. Now suppose that for some $\omega \leqq \Omega$ the following statements hold for all $\lambda < \omega$:

$(1)_\lambda$ c'_α is defined and $\leqq b_1$, and $c'_\alpha \sim a'_\alpha$ for all $\alpha < \lambda$, and

$(2)_\lambda$ $(a'_\alpha, \alpha < \Omega, c'_\beta, \beta < \lambda)\perp$.

If ω is a limit ordinal, it is clear that $(1)_\omega$ and, by Theorem 2.3, $(2)_\omega$ also hold. Suppose now that ω has a predecessor, say $\omega = \gamma + 1$. Define

$$d = [b_{\gamma+1} - \{\sum_{\alpha>\gamma} \oplus a'_\alpha \oplus (\sum_{\alpha\leqq\gamma} \oplus a'_\alpha \oplus \sum_{\alpha<\gamma} \oplus c'_\alpha)b_{\gamma+1}\}].$$

A suitable application of Theorem 2.7 shows that

$$(\sum_{\alpha\leqq\gamma} a'_\alpha)b_{\gamma+1} = (\sum_{\alpha\leqq\gamma} a'_\alpha)a_{\gamma+1}b_{\gamma+1} = 0;$$

hence, by Theorem 3.1, $(\sum_{\alpha\leqq\gamma} a'_\alpha + \sum_{\alpha<\gamma} c'_\alpha)b_{\gamma+1} \propto \sum_{\alpha<\gamma} c'_\alpha$. By Theorem 2.8, $\sum_{\alpha<\gamma} \oplus c'_\alpha \sim \sum_{\alpha<\gamma} \oplus a'_\alpha$. Hence, from the transitivity of perspectivity and Theorem 3.3,

$$(\sum_{\alpha\leqq\gamma} \oplus a'_\alpha \oplus \sum_{\alpha<\gamma} \oplus c'_\alpha)b_{\gamma+1} \oplus \sum_{\alpha>\gamma} \oplus a'_\alpha \propto \sum_{\alpha<\gamma} \oplus a'_\alpha \oplus \sum_{\alpha>\gamma} \oplus a'_\alpha.$$

Since $\sum_{\alpha<\gamma} \oplus a'_\alpha \oplus \sum_{\alpha>\gamma} \oplus a'_\alpha \oplus a'_\gamma = a \propto b_{\gamma+1}$, Theorem 3.2 implies that $a'_\gamma \propto d$. Hence we can define c'_γ with $c'_\gamma \leqq d \leqq b_{\gamma+1} \leqq b_1$, and $a'_\gamma \sim c'_\gamma$. Thus $(1)_\omega$ will hold. Furthermore,

$$\begin{aligned} c'_\gamma(\sum_{\alpha<\Omega} a'_\alpha + \sum_{\alpha<\gamma} c'_\alpha) &= c'_\gamma\, db_{\gamma+1}(\sum_{\alpha<\Omega} a'_\alpha + \sum_{\alpha<\gamma} c'_\alpha) \\ &= c'_\gamma\, d\{\sum_{\alpha>\gamma} a'_\alpha + (\sum_{\alpha\leqq\gamma} a'_\alpha + \sum_{\alpha<\gamma} c'_\alpha)b_{\gamma+1}\} = 0, \end{aligned}$$

implying (use $(2)_\gamma$ and Theorem 2.6) that $(a'_\alpha, \alpha < \Omega, c'_\beta, \beta < \omega)\perp$. Thus $(2)_\omega$ also holds. Thus, by induction on ω, we can define c'_α for all $\alpha < \Omega$ so that $(1)_\Omega$ and $(2)_\Omega$ hold.

If we now set $c = \sum_{\alpha<\Omega} \oplus c'_\alpha$, we have $a \oplus c \leqq b_1$, hence $a \oplus c \propto b_1$, and also, by Theorem 2.8, $a \sim c$. Thus, under the hypotheses of Theorem 3.5, there exists an element c such that $a \sim c$ and $a \oplus c \propto b_1$.

4. If, for any fixed α, we replace all $b_\beta, \beta \leqq \alpha$, by b_α in the preceding paragraphs, it follows that there exists an element c_α such that $a \sim c_\alpha$ and $a \oplus c_\alpha \propto b_\alpha$. From the transitivity of perspectivity and Theorem 3.3 we derive $c \sim c_\alpha$, $a \oplus c \sim a \oplus c_\alpha$, and hence $a \oplus c \propto b_\alpha$ for all $\alpha < \Omega$.

5. If we now apply the reasoning of the preceding paragraphs to $a \oplus c$ in place of a and iterate this process, we obtain an independent infinite sequence of elements, each perspective to a. By Theorem 2.12 this implies that $a = 0$, and the theorem is proved.

THEOREM 3.6. *If* $a \propto b_\alpha$ *and* $b_\alpha \geqq b_\beta$ *for all* $1 \leqq \alpha \leqq \beta < \Omega$, *then* $a \propto \prod_{\alpha<\Omega} b_\alpha$.

Proof. 1. As in the proof of Theorem 3.5 we may suppose that Ω is a limit ordinal and that $b_\gamma = \prod_{\alpha<\gamma} b_\alpha$ for all limit ordinals $\gamma \leqq \Omega$. Define $b'_\alpha = [b_\alpha - b_{\alpha+1}]$. Then, by Theorem 2.7, $b_\alpha = \sum_{\alpha \leqq \beta < \Omega} \oplus\, b'_\beta \oplus b_\Omega$ for all $\alpha \leqq \Omega$.

2. Set $c = ab_\Omega$ and define $a_1 = [a - c]$, $b'_\Omega = [b_\Omega - c]$. Then $(a_1, b'_\Omega, c)\perp$. Now suppose that for some $\omega < \Omega$ the following statements hold for all $\lambda < \omega$:

$(3)_\lambda$ a_λ, $a_{\lambda+1}$, b_Ω^λ, $b_\Omega^{\lambda+1}$, a'_λ, $b_\Omega^{\lambda\prime}$ are defined and $a'_\lambda \sim b_\Omega^{\lambda\prime}$.

$(4)_\lambda$ $a_\lambda = a_{\lambda+1} \oplus a'_\lambda$, $b_\Omega^\lambda = b_\Omega^{\lambda+1} \oplus b_\Omega^{\lambda\prime}$.

$(5)_\lambda$ $a_\lambda = \prod_{\alpha<\lambda} a_\alpha$, $b_\Omega^\lambda = \prod_{\alpha<\lambda} b_\Omega^\alpha$ if λ is a limit ordinal.

$(6)_\lambda$ $a_{\lambda+1} \propto \sum_{\lambda \leqq \alpha < \Omega} b'_\alpha$.

Then, as we shall show, a_ω, $a_{\omega+1}$, b_Ω^ω, $b_\Omega^{\omega+1}$, a'_ω, $b_\Omega^{\omega\prime}$ can be defined in such a way that $(3)_\omega$–$(6)_\omega$ also hold.

If ω has a predecessor, say $\omega = \gamma + 1$, then a_ω, b_Ω^ω are defined by $(3)_\gamma$; if ω is a limit ordinal, define $a_\omega = \prod_{\alpha<\omega} a_\alpha$, $b_\Omega^\omega = \prod_{\alpha<\omega} b_\Omega^\alpha$. It is clear that $(5)_\omega$ will hold. Whether ω is a limit ordinal or not, we have by Theorem 2.7,

$$a = a_1 \oplus c = a_\omega \oplus \sum_{\alpha<\omega} a'_\alpha \oplus c \propto b_{\omega+1} = \sum_{\omega<\alpha<\Omega} b'_\alpha \oplus b_\Omega^\omega \oplus \sum_{\alpha<\omega} b_\Omega^{\alpha\prime} \oplus c.$$

Since $(c, a_1, b'_\Omega)\perp$ and $a'_\alpha \leqq a_1$ and $b_\Omega^{\alpha\prime} \leqq b_\Omega^1$ for all α, it follows by Theorems 2.1 and 2.6 that $(c, a'_\alpha, \alpha < \omega, b_\Omega^{\beta\prime}, \beta < \omega)\perp$. Hence by $(3)_\alpha$, $\alpha < \omega$, and Theorems 2.8 and 3.3

$$\sum_{\alpha<\omega} \oplus\, a'_\alpha \oplus c \sim \sum_{\alpha<\omega} \oplus\, b_\Omega^{\alpha\prime} \oplus c.$$

Then by Theorem 3.2,

$$a_\omega \propto \sum_{\omega<\alpha<\Omega} b'_\alpha \oplus b_\Omega^\omega.$$

By Theorem 3.4 we can define $a_{\omega+1}$, a'_ω, $b_\Omega^{\omega\prime}$ such that

$$a_\omega = a_{\omega+1} \oplus a'_\omega, \qquad a_{\omega+1} \propto \sum_{\omega<\alpha<\Omega} b'_\alpha, \qquad a'_\omega \sim b_\Omega^{\omega\prime} \leqq b_\Omega^\omega.$$

If we define $b_\Omega^{\omega+1} = [b_\Omega^\omega - b_\Omega^{\omega\prime}]$, it follows that all $(3)_\omega$–$(6)_\omega$ will be satisfied.

By induction on ω, we can suppose that $(3)_\alpha$–$(6)_\alpha$ hold for all $\alpha < \Omega$. Then we can define a_Ω, b_Ω^Ω to satisfy $(5)_\Omega$.

3. From $(6)_\alpha$, $a_\Omega \leqq a_{\alpha+1} \propto \sum_{\alpha \leqq \beta < \Omega} b'_\beta$ for all $\alpha < \Omega$; by Theorem 2.4, $\prod_{\alpha<\Omega} \left(\sum_{\alpha \leqq \beta < \Omega} b'_\beta \right) = 0$ since $(b'_\beta, \beta < \Omega)\perp$. Theorem 3.5 now shows that $a_\Omega = 0$. By Theorems 2.1 and 2.6, $(c, a'_\alpha, \alpha < \Omega, b_\Omega^{\beta\prime}, \beta < \Omega)\perp$. Finally, from Theorems 2.7, 2.8, 3.3,

$$a = \sum_{\alpha<\Omega} \oplus\, a'_\alpha \oplus c \sim \sum_{\alpha<\Omega} \oplus\, b_\Omega^{\alpha\prime} \oplus c \leqq b_\Omega = \prod_{\alpha<\Omega} b_\alpha.$$

Thus $a \propto \prod_{\alpha<\Omega} b_\alpha$, as required.

THEOREM 3.7. *If* $b_\alpha \propto a$ *and* $b_\alpha \leqq b_\beta$ *for all* $1 \leqq \alpha \leqq \beta < \Omega$, *then* $\sum_{\alpha<\Omega} b_\alpha \propto a$.

Proof. By Theorem 2.14 this is equivalent to the dual theorem to Theorem 3.6.

THEOREM 3.8 (Continuity of perspectivity). *If* $\lim_{\alpha\to\Omega} a_\alpha = a$ *and* $\lim_{\alpha\to\Omega} b_\alpha = b$ *and* $a_\alpha \sim b_\alpha$ *for all* $\alpha < \Omega$, *then* $a \sim b$.

Proof. If $\beta_1, \beta_2 < \Omega$, and $\beta = \max(\beta_1, \beta_2)$, then

$$\prod_{\alpha \geqq \beta_1} a_\alpha \leqq a_\beta \sim b_\beta \leqq \sum_{\alpha \geqq \beta_2} b_\alpha .$$

Hence, by Theorem 3.6, $\prod_{\alpha \geqq \beta_1} a_\alpha \propto \prod_{\beta_2<\Omega} \sum_{\alpha \geqq \beta_2} b_\alpha = b$. Then, by Theorem 3.7, $a = \sum_{\beta_1<\Omega} \prod_{\alpha\geqq\beta_1} a_\alpha \propto b$, that is, $a \propto b$. Similarly, $b \propto a$. By Theorem 2.13 this implies $a \sim b$, and the theorem is proved.

THEOREM 3.9 (Additivity of perspectivity). *If, for any ordinal* Ω, $(a_\alpha, \alpha < \Omega)\perp$ *and* $(b_\alpha, \alpha < \Omega)\perp$ *and* $a_\alpha \sim b_\alpha$ *for all* $\alpha < \Omega$, *then* $\sum_{\alpha<\Omega} \oplus a_\alpha \sim \sum_{\alpha<\Omega} \oplus b_\alpha$.

Proof. Let $A_\alpha = \sum_{\beta<\alpha} a_\beta$, $B_\alpha = \sum_{\beta<\alpha} b_\beta$ for all $\alpha \leqq \Omega$. It is clear that $A_\alpha \leqq A_\beta$ and $B_\alpha \leqq B_\beta$ for all $\alpha \leqq \beta \leqq \Omega$ and that $A_\gamma = \lim_{\alpha\to\gamma} A_\alpha$, $B_\gamma = \lim_{\alpha\to\gamma} B_\alpha$ for all limit ordinals $\gamma \leqq \Omega$. Suppose now that we have already established $A_\alpha \sim B_\alpha$ for all $\alpha < \omega$ for some $\omega \leqq \Omega$. Then if ω is a limit ordinal, we have also, by Theorem 3.8, that $A_\omega \sim B_\omega$. If, however, ω has a predecessor λ, then $A_\lambda \sim B_\lambda$, $a_\lambda \sim b_\lambda$ together imply, by Theorem 3.3, that $A_\lambda \oplus a_\lambda \sim B_\lambda \oplus b_\lambda$, that is, that $A_\omega \sim B_\omega$. Since $A_1 = B_1 = 0$, hence $A_1 \sim B_1$, it follows by induction on ω that $A_\Omega \sim B_\Omega$, that is, that $\sum_{\alpha<\Omega} \oplus a_\alpha \sim \sum_{\alpha<\Omega} \oplus b_\alpha$. This proves the theorem.

4. **Weakening of the axioms.** Consider a system L satisfying Axioms I–V as given in §1, but with Axioms II, III restricted to sets of indices of cardinal power $< \aleph$, for some fixed $\aleph > \aleph_0$. Call such a system an $\aleph$-geometry. For any pair of elements u, v with $u \leqq v$ let $L(u, v)$ denote the subsystem of L of all elements x with $u \leqq x \leqq v$. Then if L is an $\aleph$-geometry, it is easily seen that $L(u, v)$ is also an $\aleph$-geometry. We shall call a set of elements a_α, $\alpha \in I$, independent over u, if for some $v \geqq u$, $(a_\alpha, \alpha \in I)\perp$ in $L(u, v)$.

If L is an $\aleph$-geometry and we restrict ourselves to some fixed $L(u, v)$, then all those parts of the preceding sections which involve sums and intersections of sets of elements of cardinal power $< \aleph$ will hold exactly as before.[11] And if I is a set of indices of cardinal power $< \aleph$, then every set of elements a_α, $\alpha \in I$, b_β,

[11] The transitivity of perspectivity holds since $\aleph > \aleph_0$. See [1].

$\beta \in I$, has a sum and an intersection, and hence is contained in an $L(u, v)$. Hence we can deduce the following theorems:

THEOREM 4.1. *Let L be an $\aleph$-geometry and let Ω be any limit ordinal such that the set of ordinals $\alpha < \Omega$ has cardinal power $< \aleph$. If $\lim_{\alpha \to \Omega} a_\alpha = a$ and $\lim_{\alpha \to \Omega} b_\alpha = b$, and if $a_\alpha \sim b_\alpha$ for all $\alpha < \Omega$, then $a \sim b$.*

THEOREM 4.2. *Let L be an $\aleph$-geometry and let I be any set of indices of cardinal power $< \aleph$. If a_α, $\alpha \in I$, and b_β, $\beta \in I$, are independent over the same u and $a_\alpha \sim b_\alpha$ for all $\alpha \in I$, then $\sum_{\alpha \in I} a_\alpha \sim \sum_{\alpha \in I} b_\alpha$.*[12]

BIBLIOGRAPHY

1. ISRAEL HALPERIN, Transactions of the American Mathematical Society, vol. 44(1938), pp. 537-562.
2. ISRAEL HALPERIN, Annals of Mathematics, vol. 40(1939), pp. 581-599.
3. J. VON NEUMANN, Proceedings of the National Academy of Sciences, vol. 22(1936), pp. 92-100.
4. J. VON NEUMANN, *Continuous Geometry*, planographed, Institute for Advanced Study, Princeton, New Jersey, 1936-1937.
5. J. VON NEUMANN, *Continuous Geometry*, American Mathematical Society Colloquium Lectures, to appear in book form.

HARVARD UNIVERSITY.

[12] The well-ordering theorem is used here.

THE VARIATION OF THE SIGN OF V FOR AN ANALYTIC FUNCTION $U + iV$

By M. S. Robertson

1. **Introduction.** The following theorem is M. Cartwright's [2]:[1]

Theorem A. *Let*

$$f(z) = \sum_0^\infty a_n z^n$$

be regular and multivalent of order p for $|z| < 1$ and have q zeros within this circle. Then for $r < 1$

$$|f(re^{i\theta})| < A(p)\mu_q(1 - r)^{-2p},$$

where $A(p)$ is a constant depending only upon p, and where

$$\mu_q = \max\{|a_0|, |a_1|, |a_2|, \cdots, |a_q|\}.$$

Recently, with the help of the preceding theorem, M. Biernacki [1] established the inequality for the coefficients of $f(z)$ given by the following theorem.

Theorem B. *With the same hypotheses as in Theorem* A, *the coefficients of $f(z)$ satisfy the inequality*

$$|a_n| < A(p)\mu_q n^{2p-1} \qquad (n > 0),$$

where

$$\mu_q = \max\{|a_1|, |a_2|, \cdots, |a_q|\}.$$

The following new theorems established in this paper appear to give somewhat similar inequalities under a different hypothesis. The theorems to follow below overlap Theorems A and B in some cases, especially when $f(z)$ is real on the real axis.

Let $f(re^{i\theta}) = U(r, \theta) + iV(r, \theta)$. Let $z = re^{i\theta}$ traverse the circle $|z| = r$ once, starting at any point $z_0 = re^{i\theta_0}$ where $V(r, \theta_0) \neq 0$. If $r < 1$, $V(r, \theta)$ is a continuous function of θ. As z makes a complete revolution around the circle from z_0 and back again to z_0, $V(r, \theta)$ has either (i) a constant sign or (ii) changes sign an even number of times, if we assume that the number of changes in sign is finite. We now state the theorems to be demonstrated in this paper.

Theorem 1. *Let*

$$f(z) = \sum_0^\infty a_n z^n = U + iV$$

Received December 15, 1938.

[1] Numbers in brackets refer to the list of references at the end of the paper.

be regular for $|z| < 1$. *If there is an interval* $0 < 1 - \delta < r < 1$ *such that* V *does not change sign more than* $2p$ *times on* $|z| = r$ *for any value* r *of the given interval, then*

(i) $$|a_n| < A(p)\mu_p n^{2p} \qquad (n > 0,\ p \geqq 0),$$

(ii) $$\limsup_{n\to\infty} \left| \frac{a_n}{n^{2p}} \right| \leqq 2 \sum_{k=0}^{p} \frac{|a_k|}{(p+k)!\,(p-k)!},$$

(iii) $$|f(re^{i\theta})| < A(p)\mu_p(1-r)^{-2p-1} \qquad (r < 1),$$

(iv) $$\frac{1}{2\pi}\int_0^{2\pi} |f(re^{i\theta})|\,d\theta < A(p)\mu_p(1-r)^{-2p} \qquad (p \geqq 1,\ r < 1),$$

where $A(p)$ *is a constant depending only upon* p, *and where*

$$\mu_p = \max\{|a_0|, |a_1|, |a_2|, \cdots, |a_p|\}.$$

THEOREM 2. *If, in addition to satisfying the hypothesis of Theorem* 1, $f(z)$ *is real on the real axis, then*

(i) $$|a_n| < A(p)\mu_p n^{2p-1} \qquad (n > 0,\ p \geqq 1),$$

(ii) $$|f(re^{i\theta})| < A(p)\mu_p(1-r)^{-2p} \qquad (r < 1,\ p \geqq 1),$$

(iii) $$\frac{1}{2\pi}\int_0^{2\pi} |f(re^{i\theta})|\,d\theta < A(p)\mu_p(1-r)^{1-2p} \qquad (r < 1,\ p \geqq 1).$$

The function

$$\frac{z^p + \epsilon z^{p+1}}{(1-\epsilon z)^{2p+1}}, \qquad \epsilon = \pm e^{\frac{1}{2}i\pi/p},$$

shows that in Theorem 1 the exponents of n and $(1 - r)$ cannot be replaced by smaller ones. In the same way the function $z^p(1 - z)^{-2p}$ fulfills a similar purpose for Theorem 2.

The author [3] previously established Theorems 1 and 2 in part only for the special case $\mu_{p-1} = 0$, i.e., when

$$a_0 = a_1 = a_2 = \cdots = a_{p-1} = 0.$$

In this special case the method consisted in expressing $f(z)$ in terms of a function $F(z)$ regular and having a positive real part throughout the interior of the whole unit circle. A modification of this method of proof has been adopted in this paper to make the theorems completely general. It was found necessary to replace the regular function $F(z)$ by another having a pole of order not exceeding p at the origin and which no longer has the property that its real part is positive throughout the interior of the whole circle. Instead, $F(z)$ is the limit of functions with a pole at the origin, regular on $|z| = 1$ and with positive real part on $|z| = 1$. This representation for $f(z)$ is perhaps not as interesting in itself as in the previous special case. However, it is sufficient to give the desired conclusion of the theorems above, though the results are not as sharp, as far as the constants $A(p)$, μ_p are concerned, as in the special case $\mu_{p-1} = 0$.

2. **Functions with positive real part and a pole at the origin.** Before proceeding to prove Theorem 1, we need first the following lemma.

LEMMA. *Let*

$$F(z) = \sum_{k=-p}^{\infty} A_k z^k \tag{2.1}$$

be regular for $0 < |z| < 1$ *with a pole of order not exceeding* p *at the origin. If* $\Re F(re^{i\theta}) > 0$ *for* $0 < 1 - \delta < r < 1$, *then*

$$|A_n + \bar{A}_{-n}| \leqq 2\Re A_0, \tag{2.2}$$

where $\bar{A}_{-n}$ *denotes the complex conjugate of* A_{-n}.

Proof.

$$A_n r^n = \frac{1}{2\pi}\int_0^{2\pi} F(re^{i\theta})e^{-ni\theta}\, d\theta \qquad (r < 1),$$

$$\bar{A}_{-n} r^{-n} = \frac{1}{2\pi}\int_0^{2\pi} \overline{F(re^{i\theta})}e^{-ni\theta}\, d\theta,$$

$$A_n r^n + \bar{A}_{-n} r^{-n} = \frac{1}{\pi}\int_0^{2\pi} \{\Re F(re^{i\theta})\}e^{-ni\theta}\, d\theta,$$

$$|A_n r^n + \bar{A}_{-n} r^{-n}| \leqq \frac{1}{\pi}\int_0^{2\pi} \{\Re F(re^{i\theta})\}\, d\theta = 2\Re A_0,$$

for $1 - \delta < r < 1$ since $\Re F(re^{i\theta}) > 0$ in this interval. Let $r \to 1$. Then (2.2) follows.

COROLLARY. *If* $F(z)$ *is regular for* $0 < |z| \leqq 1$, *and if merely* $\Re F(e^{i\theta}) \geqq 0$ *for all real* θ, *then* (2.2) *holds again.*

3. **Proof of Theorem** 1. We may assume that $p \geqq 1$ in the proof of Theorem 1. For if $p = 0$, $V(r, \theta)$ is of constant sign within the unit circle. Then either $if(z)$ or $-if(z)$ has a non-negative real part for $|z| < 1$. For this case the theorem is well known.

It will be sufficient to prove Theorem 1 in the case where

$$g(z) = \sum_0^{\infty} a_n z^n \tag{3.1}$$

is regular within and on the circle $|z| = 1$, assuming that the imaginary part of $g(e^{i\theta})$, or $v(\theta)$, changes sign $2p$ times on the circle $|z| = 1$. If then the conclusion to Theorem 1 is true for $g(z)$, we may let $f(tz) = g(z)$, $1 - \delta < t < 1$, where $f(z)$ satisfies the conditions of Theorem 1. Then on $|z| = 1$ the imaginary part of $f(tz)$ does not change sign more than $2p$ times. On letting $t \to 1$, we shall have Theorem 1.

Since $v(\theta)$ changes sign $2p$ times on the unit circle and is continuous, $v(\theta) = 0$

for at least $2p$ distinct values of θ. Call these values θ_i $(i = 1, 2, \cdots, 2p)$ defined so that

$$
(3.2)\qquad \begin{aligned} &0 \leqq \theta_1 < 2\pi, \qquad 0 \leqq \theta_1 < \theta_2 < \theta_3 < \cdots < \theta_{2p}, \\ &0 < \theta_i - \theta_j < 2\pi \text{ for } i > j \qquad (i, j = 1, 2, \cdots, 2p). \end{aligned}
$$

Then with a proper choice of θ_1,

$$
(3.3)\qquad \begin{aligned} v(\theta) &\geqq 0 \quad \text{for} \quad \theta_{2s-1} < \theta < \theta_{2s} \qquad (s = 1, 2, \cdots, p), \\ v(\theta) &\leqq 0 \quad \text{for} \quad \theta_{2s} < \theta < \theta_{2s+1} \qquad (s = 1, 2, \cdots, p), \end{aligned}
$$

where $\theta_{2p+1} = \theta_1 + 2\pi$. Let

$$
(3.4)\qquad \begin{aligned} &\alpha_s = \pi - \tfrac{1}{2}(\theta_{2p-s+2} - \theta_{s+1}) && (s = 2, 3, \cdots, p), \\ &\mu_s = \tfrac{1}{2}(\theta_{2p-s+2} + \theta_{s+1}), \qquad \mu_1 = \tfrac{1}{2}(\theta_1 + \theta_2) && (s = 2, 3, \cdots, p), \\ &\nu_s = \tfrac{1}{2}(\theta_{2p-s+2} - \theta_{s+1}), \qquad \nu_1 = \tfrac{1}{2}(\theta_2 - \theta_1) && (s = 2, 3, \cdots, p), \\ &\phi_1 = \tfrac{1}{2}\pi - \mu_1, \\ &\phi_s = \tfrac{3}{2}\pi - \phi_1 - \phi_2 - \cdots - \phi_{s-1} - \tfrac{1}{2}(\theta_{2p-s+2} + \theta_{s+1}) && (s = 2, 3, \cdots, p). \end{aligned}
$$

Let

$$
(3.5)\qquad \begin{aligned} h_\nu(z) &\equiv z(1 - 2z\cos\nu + z^2)^{-1}, \\ H_1(z) &\equiv \frac{g(ze^{-i\phi_1})}{h_{\nu_1}(-iz)}, \\ H_j(z) &\equiv \frac{H_{j-1}(ze^{-i\phi_j})}{h_{\alpha_j}(-iz)} \qquad (j = 2, 3, \cdots, p). \end{aligned}
$$

Each of the functions H_j $(j = 1, 2, 3, \cdots, p)$ is regular for $0 < |z| \leqq 1$. Either they each have a pole at the origin whose order does not exceed p, or they are regular at the origin.

$$
(3.6)\qquad \Im H_1(e^{i\theta}) = 2v(\theta - \phi_1)\cdot\{\sin\theta - \sin(\tfrac{1}{2}\pi - \nu_1)\}.
$$

The second factor of this product is positive for $\frac{1}{2}\pi - \nu_1 < \theta < \frac{1}{2}\pi + \nu_1$ and it is negative for $\frac{1}{2}\pi + \nu_1 < \theta < \frac{5}{2}\pi - \nu_1$. On the other hand, the first factor $v(\theta - \phi_1)$ is positive for $\frac{1}{2}\pi - \nu_1 < \theta < \frac{1}{2}\pi + \nu_1$ as well as for $p - 1$ other non-overlapping intervals wherein the second factor of the product is negative. Thus the imaginary part of $H_1(z)$ changes sign $2p - 2$ times on $|z| = 1$. A similar argument shows that $\Im H_2(z)$ changes sign $2p - 4$ times on $|z| = 1$. Finally, $\Im H_p(z)$ does not change sign at all on $|z| = 1$, and is indeed positive. We thus have

$$
(3.7)\qquad \begin{aligned} g(z) &= (-1)^{p-1} H_p(-ize^{-i\mu_p}) \cdot \prod_{j=1}^{p} h_{\nu_j}(ze^{-i\mu_j}) \\ &= i(-1)^{p-1} e^{-i\phi_p} F(z) \cdot z^p \prod_{s=1}^{2p} (1 - ze^{-i\theta_s})^{-1}, \end{aligned}
$$

where

$$\sigma_p = \tfrac{1}{2}\sum_{1}^{2p}\theta_j, \qquad F(z) = -iH_p(-ize^{-i\sigma_p}),$$
(3.8)
$$F(z) = \sum_{k=-p}^{\infty} A_k z^k.$$

$F(z)$ is regular for $0 < |z| \leqq 1$ and $\Re F(e^{i\theta}) \geqq 0$. $F(z)$ is either regular at $z = 0$ or has a pole of order not exceeding p, depending upon whether some of the first coefficients $a_0, a_1, \cdots, a_p$ of $g(z)$ are different from zero or not.

From (3.1), (3.7) and (3.8) we have, on letting

(3.9) $$\prod_{s=1}^{2p}(1 - ze^{-i\theta_s}) = \sum_{s=0}^{2p} c_s z^s,$$

the identity

(3.10) $$\sum_{s=0}^{2p} c_s z^s \cdot \sum_{m=0}^{\infty} a_m z^m = i(-1)^{p-1}e^{-i\sigma_p}\cdot\sum_{s=0}^{\infty} A_{s-p} z^s,$$

whence we have, on equating coefficients of z^n from either side of (3.10),

(3.11) $$A_{n-p} = i(-1)^p e^{i\sigma_p}\cdot\sum_{k=0}^{n} a_k c_{n-k}.$$

Since $(1 + z)^{2p}$ dominates $\prod_{s=1}^{2p}(1 - ze^{-i\theta_s})$, we have

(3.12) $$|c_n| \leqq \frac{(2p)!}{(2p-n)!\,n!} \qquad (n = 0, 1, 2, \cdots, 2p).$$

Using (3.12), we obtain from (3.11)

(3.13) $$|A_{n-p}| \leqq (2p)!\sum_{k=0}^{n}\frac{|a_k|}{(k+2p-n)!(n-k)!}.$$

Let $\mu_p = \max\{|a_0|, |a_1|, |a_2|, \cdots, |a_p|\}$. Then for $n = 0, 1, 2, \cdots, p$ we have

$$|A_{-n}| \leqq \mu_p\sum_{s=0}^{p-n}\frac{(2p)!}{(2p-s)!\,s!}$$
(3.14)
$$\leqq \tfrac{1}{2}\mu_p\left[\sum_{s=0}^{2p}\frac{(2p)!}{(2p-1)!\,s!} + \frac{(2p)!}{(p!)^2}\right]$$
$$\leqq \mu_p\left[2^{2p-1} + \frac{(2p)!}{2(p!)^2}\right].$$

From (2.2) we have for $n > p$

(3.15) $$|A_n| \leqq 2|A_0| \leqq \mu_p\left[2^{2p} + \frac{(2p)!}{(p!)^2}\right],$$

while for $0 < n \leqq p$

$$|A_n| \leqq 2|A_0| + |A_{-n}| \leqq \tfrac{3}{2}\mu_p\left[2^{2p} + \frac{(2p)!}{(p!)^2}\right]. \tag{3.16}$$

From (3.7) it follows that $g(z)$ is dominated by

$$(1 - z)^{-2p} \cdot \sum_{s=0}^{\infty} |A_{s-p}| z^s = \sum_{q=0}^{\infty} \frac{(2p + q - 1)!}{(2p - 1)!q!} z^q \cdot \sum_{s=0}^{\infty} |A_{s-p}| z^s.$$

Consequently,

$$\begin{aligned} |a_n| &\leqq \sum_{q=0}^{n} \frac{(2p + q - 1)!}{(2p - 1)!q!} |A_{n-p-q}| \\ &\leqq \tfrac{3}{2}\mu_p\left[2^{2p} + \frac{(2p)!}{(p!)^2}\right] \cdot \sum_{q=0}^{n} \frac{(2p + q - 1)!}{(2p - 1)!q!} \\ &< A(p)\mu_p n^{2p} \qquad\qquad (n > 0). \end{aligned} \tag{3.17}$$

Again, from (3.13) and (3.15) we have for $n > p$

$$|A_n| \leqq 2|A_0| \leqq 2(2p)! \sum_{k=0}^{p} \frac{|a_k|}{(p + k)!(p - k)!}, \tag{3.18}$$

and $g(z)$ is dominated by

$$(1 - z)^{-2p} \cdot \sum_{s=0}^{2p} |A_{s-p}| z^s + \frac{z^{2p+1}}{(1 - z)^{2p+1}}\left[2(2p)! \sum_{k=0}^{p} \frac{|a_k|}{(p + k)!(p - k)!}\right],$$

whence

$$\limsup_{n\to\infty} \left|\frac{a_n}{n^{2p}}\right| \leqq 2 \sum_{k=0}^{p} \frac{|a_k|}{(p + k)!(p - k)!}. \tag{3.19}$$

This completes the proof of the first two parts of Theorem 1. The third part follows immediately from the first part, since for $r < 1$

$$\begin{aligned} |f(re^{i\theta})| &\leqq \sum_{n=0}^{\infty} |a_n| r^n < A(p)\mu_p(1 + \sum_{n=1}^{\infty} n^{2p} r^n) \\ &< A(p)\mu_p(1 - r)^{-2p-1}. \end{aligned} \tag{3.20}$$

To prove the fourth part of Theorem 1, we write (3.7) as

$$g(z) = i(-1)^{p-1} e^{-i\varphi_p} F(z) \phi^p(z), \tag{3.21}$$

where

$$\phi(z) = z \prod_{s=1}^{2p} (1 - ze^{-i\theta_s})^{-1/p},$$

$$\frac{z\phi'(z)}{\phi(z)} = \frac{1}{2p} \sum_{s=1}^{2p} \left(\frac{1 + ze^{-i\theta_s}}{1 - ze^{-i\theta_s}}\right),$$

whence $\Re z\phi'(z)/\phi(z) > 0$ for $|z| < 1$. Thus $\phi(z)$ is univalent and star-like for $|z| < 1$. It follows for $r < 1$ that [4]

$$\frac{1}{2\pi}\int_0^{2\pi}\left|\frac{\phi(re^{i\theta})}{r}\right|^p d\theta \leqq \frac{1}{2\pi}\int_0^{2\pi}\frac{d\theta}{|1-re^{i\theta}|^{2p}} < A(p)(1-r)^{1-2p} \qquad (p \geqq 1). \tag{3.22}$$

Also we have by (3.14), (3.15) and (3.16)

$$\max_{|z|=r}|z^p F(z)| \leqq \sum_{s=0}^{\infty}|A_{s-p}|r^s \leqq A(p)\mu_p(1-r)^{-1}. \tag{3.23}$$

Combining (3.22) and (3.23), we have for $p \geqq 1, r < 1$,

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi}|g(re^{i\theta})|\,d\theta &= \frac{1}{2\pi}\int_0^{2\pi}|z^pF(z)|\cdot\left|\frac{\phi(z)}{z}\right|^p d\theta\\ &\leqq \frac{A(p)\mu_p}{1-r}\cdot\frac{1}{2\pi}\int_0^{2\pi}\left|\frac{\phi(re^{i\theta})}{r}\right|^p d\theta\\ &\leqq A(p)\mu_p(1-r)^{-2p},\end{aligned} \tag{3.24}$$

where $A(p)$ is a constant depending only upon p. This completes the proof of Theorem 1.

4. **Proof of Theorem** 2. It will be sufficient to prove Theorem 2 for

$$g(z) = \sum_0^{\infty} a_n z^n, \qquad a_n \text{ real}, \tag{4.1}$$

when $g(z)$ is regular for $|z| \leqq 1$, $g(z)$ real on the real axis, and when the imaginary part of $g(e^{i\theta})$ changes sign $2p$ times on the unit circle. Since the coefficients are real, we have (if not for $g(z)$ then for $-g(z)$)

$$\begin{aligned}&\theta_1 = 0, \qquad \theta_{p+1} = \pi, \qquad \theta_{2p-s+2} = 2\pi - \theta_s \qquad (s = 2, 3, \cdots, p),\\ &\sigma_p = \tfrac{1}{2}\sum_1^{2p}\theta_s = \tfrac{1}{2}\theta_2 + \sum_2^{p}[\pi + \tfrac{1}{2}(\theta_{s+1} - \theta_s)] = \tfrac{1}{2}(2p-1)\pi.\end{aligned} \tag{4.2}$$

From (3.7) we then have

$$g(z) = \frac{F(z)}{1-z^2}\cdot\frac{z^p}{\prod_{j=2}^{p}(1 - 2z\cos\theta_j + z^2)}. \tag{4.3}$$

Thus $g(z)$ is dominated by the function

$$(1-z^2)^{-1}(1-z)^{2-2p}\cdot\sum_{s=0}^{\infty}|A_{s-p}|z^s \tag{4.4}$$

which on account of (3.14), (3.15) and (3.16) is in turn dominated by

$$(1-z^2)^{-1}(1-z)^{2-2p}\cdot A(p)\mu_p\left(\frac{1+z}{1-z}\right) = \frac{A(p)\mu_p}{(1-z)^{2p}}. \tag{4.5}$$

Thus

$$|a_n| < A(p)\mu_p n^{2p-1} \qquad (n > 0,\ a_n \text{ real}). \tag{4.6}$$

$$|f(re^{i\theta})| < \frac{A(p)\mu_p}{(1-r)^{2p}}. \tag{4.7}$$

To prove the third part of Theorem 2 let

$$\psi(z) = \frac{z}{\left[\prod_{j=2}^{p} (1 - 2z\cos\nu_j + z^2)\right]^{1/(p-1)}} \qquad (p > 1), \tag{4.8}$$

$$\Re\frac{z\psi'(z)}{\psi(z)} = \frac{1}{p-1}\sum_{j=2}^{p}\Re\left\{\frac{1-z^2}{1-2z\cos\nu_j+z^2}\right\} > 0 \qquad (|z| < 1), \tag{4.9}$$

$$g(z) = \frac{z^p F(z)}{1-z^2}\cdot\left\{\frac{\psi(z)}{z}\right\}^{p-1}, \tag{4.10}$$

where $\psi(z)$ is star-like for $|z| < 1$. Hence for $z = re^{i\theta}$, $p > 1$,

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi}|g(z)|\,d\theta &< \frac{A(p)\mu_p}{1-r}\cdot\frac{1}{2\pi}\int_0^{2\pi}\left|\frac{1}{1-z^2}\right|\cdot\left|\frac{\psi(z)}{z}\right|^{p-1}d\theta\\ &< \frac{A(p)\mu_p}{1-r}\left\{\frac{1}{2\pi}\int_0^{2\pi}\frac{d\theta}{|1-z^2|^2}\right\}^{\frac12}\cdot\left\{\frac{1}{2\pi}\int_0^{2\pi}\left|\frac{\psi(z)}{z}\right|^{2p-2}d\theta\right\}^{\frac12}\\ &< \frac{A(p)\mu_p}{1-r}\left(\frac{A}{1-r}\right)^{\frac12}\left(\frac{A(p)}{(1-r)^{4p-5}}\right)^{\frac12}\\ &< A(p)\mu_p(1-r)^{1-2p} \qquad (p > 1,\ r < 1).\end{aligned} \tag{4.11}$$

We shall show that this last inequality is also true when $p = 1$. Since a_0 and a_1 are real, and $a_1 \neq 0$ for $p = 1$, the imaginary part of $[g(z) - g(0)]/a_1$ is the same as for the function $g(z)/a_1$ and is positive only when the imaginary part of z is positive. By a theorem on typically real functions [4] we then have

$$\frac{1}{2\pi}\int_0^{2\pi}|g(re^{i\theta}) - g(0)|\,d\theta \leqq \frac{r|a_1|}{1-r^2}, \tag{4.12}$$

whence (4.11) follows for $p = 1$.

References

1. M. Biernacki, *Sur les fonctions multivalentes d'ordre p*, Paris Comptes Rendus, vol. 204(1936), pp. 449–451.
2. M. Cartwright, *Some inequalities in the theory of functions*, Mathematische Annalen, vol. 111(1935), pp. 98–118.
3. M. S. Robertson, *A representation of all analytic functions in terms of functions with positive real part*, Annals of Mathematics, vol. 38(1937), pp. 770–783.
4. M. S. Robertson, *On the theory of univalent functions*, Annals of Mathematics, vol. 37(1936), pp. 374–408.

Rutgers University.

CONVEXITY IN A LINEAR SPACE WITH AN INNER PRODUCT

By David Moskovitz and L. L. Dines

1. **Introduction.** In Euclidean space, a point set $\mathfrak{M}$ is said to be *linearly connected* (or convex) if, when $x \equiv (x_1, x_2, \cdots, x_n)$ and $y \equiv (y_1, y_2, \cdots, y_n)$ are any two points of the set, all points on the segment joining x and y also belong to the set. Analytically:

(1) *If x and y are in $\mathfrak{M}$, then $kx + (1 - k)y$ is in $\mathfrak{M}$ for $0 \leqq k \leqq 1$.*

A line (in two-dimensional Euclidean space) or a hyperplane (in n-dimensional Euclidean space) is said to be a *supporting* line or hyperplane of a set $\mathfrak{M}$ if it passes through at least one point of $\mathfrak{M}$ and does not separate the points of $\mathfrak{M}$. A set $\mathfrak{M}$ such that through each of its boundary points there can be passed a supporting hyperplane is said to be *completely supported at its boundary points.*

In Euclidean space, the following theorems are well known.[1]

Theorem A. *If a point set is linearly connected, it is completely supported at its boundary points.*

Theorem B. *If a point set is closed, possesses inner points, and is completely supported at its boundary points, then it is linearly connected.*[2]

The purpose of this paper is to examine how far these theorems persist in a more general space. Obviously their validity will depend on the space and the definitions assumed.

We center our attention on a space $\mathfrak{S}$ which is *linear* and in which an inner product is defined. That is, for any two elements x and y of our space there is assumed to exist a unique real-valued number denoted by $((x, y))$ and called the *inner product* of x and y. This inner product is assumed to satisfy the following:

$$((cx, y)) = c((x, y)) \qquad \text{for every real number } c,$$

$$((x + y, z)) = ((x, z)) + ((y, z)),$$

$$((x, y)) = ((y, x)),$$

$$((x, x)) \geqq 0,$$

with the equality holding if and only if x is the zero element of our space, and

$$|((x, y))| \leqq ((x, x))^{\frac{1}{2}} \cdot ((y, y))^{\frac{1}{2}}.$$

Received January 18, 1939; presented to the American Mathematical Society, September 6, 1938.

[1] See, for example, American Mathematical Monthly, vol. 45(1938), p. 202.

[2] For a symmetrical theorem, but one which is weaker than Theorems A and B, we have the following: *If a set $\mathfrak{M}$ is closed and possesses inner points, then linear connectedness of the set $\mathfrak{M}$ implies complete support at its boundary, and conversely.*

If we denote the non-negative real number $((x, x))^{\frac{1}{2}}$ by $\| x \|$, the function $\| x - y \|$ of the two elements x, y has all the necessary properties for a distance function and serves as a metric for the space. In particular, the important triangle inequality

$$\| x + y \| \leqq \| x \| + \| y \|$$

is satisfied. Convergence in terms of this metric is defined in the usual manner, and the space $\mathfrak{S}$ thus metricized is assumed to be *complete*.

We generalize the Euclidean hyperplane in what appears to be a natural manner; namely, if α is a point of our space such that $\| \alpha \| \neq 0$, the points x which satisfy an equation of the form

$$((\alpha, x)) = \text{constant}$$

will be said to form a *plane*.[3] In particular, the plane

$$((\alpha, x)) = ((\alpha, x_0)),$$

or what is the same thing

$$((\alpha, x - x_0)) = 0,$$

will be called a plane through the point x_0.

The plane

$$\pi(x) \equiv ((\alpha, x)) - c = 0$$

determines two closed half-spaces in $\mathfrak{S}$, one consisting of those points x for which $\pi(x) \geqq 0$, the other of those for which $\pi(x) \leqq 0$. The corresponding open half-spaces are represented if the equality signs be omitted.

Any set $\mathfrak{M}$ serves to classify all planes of the space $\mathfrak{S}$ into three categories with associated names as follows. Relative to the set $\mathfrak{M}$, a plane π is called (1) a *bounding* plane if all points of $\mathfrak{M}$ lie in one of the open half-spaces determined by π, (2) a *separating* plane if points of $\mathfrak{M}$ lie in each of the open half-spaces, or (3) a *supporting* plane if all points of $\mathfrak{M}$ lie in one of the closed half-spaces and at least one point of $\mathfrak{M}$ lies on π.

Let $\mathfrak{M}$ denote a set of points in the space $\mathfrak{S}$. An element of $\mathfrak{S}$ which is not a point of $\mathfrak{M}$ will be called an *exterior* point of $\mathfrak{M}$; a point of $\mathfrak{M}$ which is the limit of exterior points will be called a *boundary* point of $\mathfrak{M}$; and a point of $\mathfrak{M}$ which is not a boundary point will be called an *inner* point of $\mathfrak{M}$.

A set $\mathfrak{M}$ will be said to be *linearly connected* if condition (1) is satisfied.

Assuming these definitions, we shall prove that Theorem B generalizes completely (Theorem 6) but Theorem A does not. An example is given in §4 which exhibits a linearly connected set which has boundary points through which no supporting plane exists.[4]

[3] For the sake of brevity, we will use the term *plane* instead of *hyperplane*.

[4] In a paper published in Annali di Matematica, vol. 10(1932), entitled *Sugli spazi lineari e loro lineare varietà*, Ascoli shows, with a different definition of a plane from the one we use, that in a separable space a linearly connected set which possesses inner points

Since Theorem A does not generalize, it becomes desirable to characterize those boundary points of a linearly connected set through which supporting planes exist. We call such boundary points *normal* points, and show that they are projections (to be defined below) of exterior points of the set $\mathfrak{M}$. If y_0 is an exterior point of the set $\mathfrak{M}$, we will call the distance from y_0 to $\mathfrak{M}$ the greatest lower bound of the aggregate of values of $\| y_0 - x \|$ as x varies over $\mathfrak{M}$. If the distance from y_0 to $\mathfrak{M}$ is attained at a point x_0 of $\mathfrak{M}$, we will call x_0 a *projection of* y_0 *on* $\mathfrak{M}$. We shall show (Theorem 11) that a linearly connected set $\mathfrak{M}$ is supported at those boundary points and only those which are the projections of exterior points of $\mathfrak{M}$; and (Theorem 14) in case the set $\mathfrak{M}$ is closed as well as linearly connected, then the set of such normal points of $\mathfrak{M}$ is everywhere dense on the boundary.

In §2, we establish some useful geometric properties of the space $\mathfrak{S}$ which are needed in the later developments. We close our work (Theorems 15, 16, and corollaries) by showing that, with certain additional assumptions on our space, Theorem A generalizes completely.

2. Some geometric properties of the space $\mathfrak{S}$.

THEOREM 1. *The projection of an exterior point y_0 upon the line of points*

$$x = x_0 + t(x_1 - x_0) \tag{2}$$

is unique and is obtained when the parameter t has the value

$$t = \frac{((y_0 - x_0, x_1 - x_0))}{\| x_1 - x_0 \|^2}. \tag{3}$$

Proof. From the definition of $\| y_0 - x \|$ it follows that

$$\| y_0 - x \|^2 = \| y_0 - x_0 \|^2 + t^2 \| x_1 - x_0 \|^2 - 2t((y_0 - x_0, x_1 - x_0)),$$

and from elementary considerations it follows that this quadratic in t has its minimum value when t has the value (3).

THEOREM 2. *The projection of the exterior point y_0 upon the plane*

$$\pi(x) \equiv ((y_0 - x_0, x - x_0)) = 0 \tag{4}$$

is the point x_0.

Proof. If x_1 is any point of the plane distinct from x_0, the projection of y_0 upon the line (2) is given by the parameter value (3). Since x_1 satisfies (4), this parameter value reduces to zero, and the projection is x_0.

(a convex body) is completely supported at its boundary. However, in this paper we do not assume our space to be separable and we do not restrict our set $\mathfrak{M}$ to be a convex body.

Since the present paper was submitted for publication, we have been able to prove without the assumption of separability that a linearly connected and closed set which contains inner points is completely supported at the boundary.

THEOREM 3. *The projection of a point y_0 on a plane*

$$\pi(x) \equiv ((\alpha, x)) - b = 0, \qquad \| \alpha \| = 1,$$

is unique and is given by

$$x_0 = y_0 + t_0\alpha, \qquad \text{where } t_0 = b - ((\alpha, y_0)).$$

Proof. It is easily verified that x_0 is on the plane π. Suppose x_1 is any other point on π; then

$$\begin{aligned} \| y_0 - x_1 \|^2 &= \| y_0 - x_0 + x_0 - x_1 \|^2 \\ &= \| y_0 - x_0 \|^2 + \| x_0 - x_1 \|^2 - 2t_0((\alpha, x_0 - x_1)) \\ &= \| y_0 - x_0 \|^2 + \| x_0 - x_1 \|^2, \end{aligned}$$

the last equality following from the fact that both x_0 and x_1 are on π. Hence

$$\| y_0 - x_1 \| > \| y_0 - x_0 \|,$$

and x_0 is the projection.

The following corollary is easily established. We give it without proof, since it is a special case of a more general theorem (Theorem 9) which is established later.

COROLLARY. *If x_0 is the projection on a plane π of the point y_0, it is also the projection on π of any point on the line $\overline{x_0y_0}$.*

THEOREM 4. *Through a given point x_0 there is one and only one plane on which x_0 is the projection of a given point y_0. Its equation may be written*

$$\pi(x) \equiv ((y_0 - x_0, x - x_0)) = 0.$$

Proof. That x_0 is the projection of y_0 on this plane was shown in Theorem 2. To prove the uniqueness of this plane, let us suppose that

$$\pi'(x) \equiv ((\alpha, x - x_0)) = 0, \qquad \| \alpha \| = 1,$$

is any plane through x_0 on which x_0 is the projection of y_0. Then by Theorem 3 it follows that

$$x_0 = y_0 + t_0\alpha,$$

where t_0 is a well-defined real number. Hence

$$\alpha = -\frac{1}{t_0}(y_0 - x_0)$$

and

$$\pi'(x) \equiv -\frac{1}{t_0}\pi(x);$$

that is, the planes $\pi'(x) = 0$ and $\pi(x) = 0$ are identical.

DEFINITION. The set of points x for which $\| x - x_0 \| \leqq r$ will be called a *sphere* of radius r about the point x_0. The set of points for which the equality holds will be the boundary points of the sphere.

It is easily verified that the sphere is a linearly connected set.

The following lemmas are needed in the proofs of Theorem 5 and subsequent theorems.

LEMMA 1. *If x_0 is an inner point of a set $\mathfrak{M}$, there is a sphere about x_0 in which every point belongs to $\mathfrak{M}$.*

Proof. Suppose this were not true; then corresponding to any preassigned monotone decreasing sequence of positive numbers $\{r_n\}$ there would be a sequence of exterior points $\{x_n\}$ such that

$$\| x_0 - x_n \| < r_n .$$

The greatest lower bound of the sequence $\{r_n\}$ cannot be zero for if it were, x_0 would be a limit point of $\{x_n\}$, in which case x_0 would be a boundary point of $\mathfrak{M}$. If the greatest lower bound of the sequence $\{r_n\}$ is the positive number r, then in a sphere of radius less than r about x_0 there are no exterior points of $\mathfrak{M}$.

LEMMA 2. *A supporting plane of a set $\mathfrak{M}$ cannot contain an inner point of $\mathfrak{M}$.*

Proof. Let

$$\pi(x) \equiv ((\alpha, x)) - b = 0, \qquad \| \alpha \| = 1,$$

be a supporting plane of $\mathfrak{M}$. Let x_0 be an inner point of $\mathfrak{M}$ which, if possible, is on the plane π, and let r be the radius of the spherical neighborhood about x_0 each point of which belongs to $\mathfrak{M}$. Then the points

$$x_1 = x_0 + r\alpha \quad \text{and} \quad x_2 = x_0 - r\alpha$$

are in $\mathfrak{M}$. But

$$\pi(x_1) = ((\alpha, x_0 + r\alpha)) - b = r,$$

and

$$\pi(x_2) = ((\alpha, x_0 - r\alpha)) - b = -r.$$

Consequently, x_1 and x_2 are separated by the plane π, and the hypothesis that π is a supporting plane of $\mathfrak{M}$ is contradicted. The contradiction proves that x_0 cannot be on a supporting plane.

THEOREM 5. *Through each boundary point of a sphere there passes one and only one supporting plane. It is that plane on which the boundary point is the projection of the center of the sphere.*

Proof. If the center of the sphere is y_0 and its radius r, then the boundary point x_0 may be expressed in the form

$$x_0 = y_0 + ru_0 , \qquad \| u_0 \| = 1. \tag{5}$$

That the plane

$$\pi(x) \equiv ((y_0 - x_0, x - x_0)) = 0$$

is actually a supporting plane may be verified as follows. Substitution for x_0 of its value from (5) in $\pi(x)$ gives

$$\pi(x) = ((-ru_0, x - y_0 - ru_0)) = r^2 - A, \tag{6}$$

where

$$A = r((u_0, x - y_0)).$$

Since

$$| A | \leqq r \| u_0 \| \cdot \| x - y_0 \| = r \| x - y_0 \|,$$

it follows from (6) and the definition of the sphere that, for all points x of the sphere, we have $\pi(x) \geqq 0$.

The uniqueness of the supporting plane follows immediately from Theorem 4 in view of the fact that x_0 must be the projection of the center y_0. If some point $x_0' \neq x_0$ were the projection of y_0, its distance from y_0 would be less than r, and so it would be an inner point of the sphere. But, by Lemma 2, an inner point of a set cannot be on a supporting plane of that set.

3. **Generalization of Theorem B.**

THEOREM 6. *If $\mathfrak{M}$ is closed, possesses inner points, and is completely supported at its boundary points, then it is linearly connected.*

Proof. Suppose $\mathfrak{M}$ is not linearly connected; then there is a segment joining two points x_1 and x_2 of $\mathfrak{M}$ on which there lies a point y exterior to $\mathfrak{M}$. We choose some interior point x_3 of $\mathfrak{M}$ and denote by v the (certainly existing) boundary point on the segment joining x_3 to y. From their descriptions, y and v may be expressed in the forms

$$y = kx_1 + (1 - k)x_2, \qquad v = k'y + (1 - k')x_3$$

where $0 < k < 1$ and $0 < k' < 1$, and from this it follows that v may be expressed in the form

$$v = a_1x_1 + a_2x_2 + a_3x_3, \tag{7}$$

where

$$a_1 > 0, \qquad a_2 > 0, \qquad a_3 > 0. \tag{8}$$

Now, by hypothesis, there exists through the boundary point v a supporting plane $\pi(x) = 0$, the definition thereof necessitating that $\pi(v) = 0$, while no two of the real constants $\pi(x_1)$, $\pi(x_2)$, $\pi(x_3)$ differ in sign. But by (7)

$$\pi(v) = a_1\pi(x_1) + a_2\pi(x_2) + a_3\pi(x_3),$$

and so, in view of (8),

$$\pi(x_1) = \pi(x_2) = \pi(x_3) = 0.$$

But x_3 is an *inner* point of $\mathfrak{M}$, and the condition $\pi(x_3) = 0$ implies that x_3 is on a supporting plane, and this was shown in Lemma 2 to be impossible.

Therefore our original supposition was incorrect and $\mathfrak{M}$ is linearly connected.

For a stronger theorem than that given in the preceding, we have the following.

THEOREM 7. *If $\mathfrak{M}$ is closed, possesses inner points, and is supported at a set of points which is everywhere dense on the boundary, then $\mathfrak{M}$ is linearly connected.*

Proof. Let x_1 and x_2 be two points of $\mathfrak{M}$ on whose segment, if possible, there is an exterior point y. Let x_3' be an inner point of $\mathfrak{M}$; then there exists a boundary point v' on the segment joining x_3' to y. Let

$$D = \| x_3' - y \| \qquad \text{and} \qquad d = \| v' - y \|.$$

Then

$$x_3' - y = \frac{D}{d}(v' - y),$$

and

$$x_3' = y + \frac{D}{d}(v' - y).$$

Let r be the radius of the neighborhood about x_3' for which all points belong to $\mathfrak{M}$. For any preassigned $\epsilon > 0$, there exists at least one normal boundary point v such that $\| v - v' \| < \epsilon$. Choose $\epsilon = rd/D$. Let

$$v = v' + c\lambda,$$

where $\| \lambda \| = 1$, and $c < rd/D$, be such a normal point. Let

$$x_3 = y + \frac{D}{d}(v - y);$$

then x_3 is on the line joining y to v but not on the segment $\overline{yv}$ since $D/d > 1$. Also

$$x_3 - x_3' = \frac{D}{d}(v - v'); \qquad \| x_3 - x_3' \| = \frac{D}{d} \| v - v' \| = \frac{D}{d} \cdot c < r.$$

Hence x_3 is an *inner* point of $\mathfrak{M}$, and the conditions upon the points x_1, x_2, x_3, y, and v are precisely those upon the similarly named points in Theorem 6, and from these followed the contradiction. The same argument therefore proves that $\mathfrak{M}$ is indeed linearly connected.

4. On the generalization of Theorem A.

THEOREM 8. *If the exterior point y has a projection on the linearly connected set $\mathfrak{M}$, this projection is unique.*

Proof. Suppose x_1 and x_2 were distinct projections of y on $\mathfrak{M}$. Then the point

$$z = \tfrac{1}{2}(x_1 + x_2)$$

would be distinct from x_1 and x_2 but would belong to $\mathfrak{M}$, and we easily obtain the impossible conclusion

$$\| y - z \| < \| y - x_1 \| . \tag{9}$$

For

$$\begin{aligned} \| y - x_1 \|^2 &= \| y - z + z - x_1 \|^2 \\ &= \| y - z \|^2 + \| z - x_1 \|^2 + 2P, \end{aligned} \tag{10}$$

where

$$\begin{aligned} P &= ((y - z, z - x_1)) \\ &= ((\tfrac{1}{2}[y - x_1] + \tfrac{1}{2}[y - x_2], \tfrac{1}{2}[y - x_1] - \tfrac{1}{2}[y - x_2])) \\ &= \tfrac{1}{4}[\, \| y - x_1 \|^2 - \| y - x_2 \|^2] = 0. \end{aligned} \tag{11}$$

From (10) and (11) follows the contradiction (9), and the contradiction proves the theorem.

DEFINITION. If x_1 is an exterior point of $\mathfrak{M}$ and x_0 is its projection, the set of points x_c of the form

$$x_c = x_0 + c(x_1 - x_0) \qquad (c > 0) \tag{12}$$

will be called the *projector* of x_1 on $\mathfrak{M}$.

THEOREM 9. *If, on a linearly connected set $\mathfrak{M}$, x_0 is the projection of an exterior point x_1, it is likewise the projection of all points on the projector of x_1.*

Proof. Let x_c, given by (12), be a point on the projector of x_1.

We first prove the theorem for the case $0 < c < 1$, by showing that the existence of a point y of $\mathfrak{M}$ for which

$$\| x_c - y \| < \| x_c - x_0 \| \tag{13}$$

implies the contradiction

$$\| x_1 - y \| < \| x_1 - x_0 \| . \tag{14}$$

By a sequence of simple reductions we have

$$\begin{aligned} \| x_1 - y \| &\leqq \| x_1 - x_c \| + \| x_c - y \| \\ &< \| x_1 - x_c \| + \| x_c - x_0 \| , && \text{by (13)} \\ &= \| x_1 - x_0 - c(x_1 - x_0) \| + \| c(x_1 - x_0) \| , && \text{by (12)} \\ &= \| x_1 - x_0 \| , \end{aligned}$$

and the contradiction (14) follows.

In case $c > 1$, from the definition of x_c, we have

$$x_1 = x_0 + c^{-1}(x_c - x_0) \qquad (0 < c^{-1} < 1).$$

If the projection of x_c on $\mathfrak{M}$ were y, then from the result of the first case the projection of x_1 on $\mathfrak{M}$ would also be y. But the projection of x_1 on $\mathfrak{M}$ is x_0, and since this projection is unique (Theorem 8), we must have $y = x_0$; hence the projection of x_c on $\mathfrak{M}$ is x_0.

THEOREM 10. *If x_0 is the projection of an exterior point y_0 upon the linearly connected set $\mathfrak{M}$, then*

$$\pi(x) \equiv ((y_0 - x_0, x - x_0)) = 0$$

is the equation of a supporting plane through x_0.

Proof. It will be sufficient to show that $\pi(x) \leqq 0$ for all points x of $\mathfrak{M}$. To this end, we suppose that there is a point x_1 of $\mathfrak{M}$ for which $\pi(x_1) > 0$ and obtain the contradictory result that x_0 is not the projection of y_0 on $\mathfrak{M}$.

Since $\mathfrak{M}$ is linearly connected, points of the line

$$x = x_0 + t(x_1 - x_0) \tag{15}$$

corresponding to parameter values $0 \leqq t \leqq 1$ belong to $\mathfrak{M}$. It is easily verified that y_0 cannot belong to the line (15) without contradicting one of the conditions $\pi(y_0) > 0$, y_0 is exterior to $\mathfrak{M}$, or $\| y_0 - x_0 \| < \| y_0 - x_1 \|$. Hence no point of the projector

$$\gamma = x_0 + \epsilon(y_0 - x_0) \qquad (\epsilon > 0) \tag{16}$$

with the exception of the projection x_0 itself lies on (15). We may therefore use formula (3) to find the projection of any point γ of (16) on the line (15). The result is

$$p_\gamma = x_0 + \epsilon A(x_1 - x_0),$$

where

$$A = \frac{((y_0 - x_0, x_1 - x_0))}{\| x_1 - x_0 \|^2} = \frac{\pi(x_1)}{\| x_1 - x_0 \|^2}.$$

Since $\pi(x_1)$ is by assumption positive, ϵ can be chosen positive and so small that p_γ belongs to the segment of (15) in $\mathfrak{M}$; we so choose ϵ. Now the distance $\| \gamma - x_0 \|$ is greater than the distance $\| \gamma - p_\gamma \|$. For

$$\gamma - x_0 = (\gamma - p_\gamma) + (p_\gamma - x_0),$$

$$\| \gamma - x_0 \|^2 = \| \gamma - p_\gamma \|^2 + \| p_\gamma - x_0 \|^2,$$

because, as may be easily verified,

$$((\gamma - p_\gamma, p_\gamma - x_0)) = \epsilon^2[A((y_0 - x_0, x_1 - x_0)) - A^2 \| x_1 - x_0 \|^2] = 0.$$

Thus, x_0 is not the projection of γ on $\mathfrak{M}$. By Theorem 9, x_0 is therefore not the projection of y_0, and the desired contradiction is reached.

THEOREM 11. *A linearly connected set $\mathfrak{M}$ has supporting planes at those, and only those, of its boundary points which are the projections of exterior points.*

Proof. If x_0 is the projection of an exterior point on $\mathfrak{M}$, the existence of a supporting plane through x_0 follows from Theorem 10.

Suppose, conversely, that x_0 is a boundary point through which passes a supporting plane

$$\pi(x) \equiv ((\alpha, x - x_0)) = 0, \qquad \| \alpha \| \neq 0;$$

and suppose for definiteness that

$$\pi(x) \leqq 0$$

for all points x of $\mathfrak{M}$.

If we define y_0 by the relation

$$y_0 = x_0 + \alpha,$$

then

$$\| y_0 - x \|^2 = \| x_0 - x + \alpha \|^2 = \| x_0 - x \|^2 + \| \alpha \|^2 - 2\pi(x)$$

which, for all x belonging to $\mathfrak{M}$, is obviously non-negative and attains its minimum value $\| \alpha \|^2$ when and only when $x = x_0$. Therefore x_0 is the projection of y_0 on $\mathfrak{M}$.

We now state and outline the proof of a theorem which will be useful in further developments, and which has perhaps some inherent interest. Some tedious algebraic details are omitted in the proof, but they are similar to others which have appeared earlier in the paper.

THEOREM 12. *If $\mathfrak{N}$ is a linearly connected set of points belonging to the "spherical shell" defined by*

$$r \leqq \| x - \alpha \| \leqq r + h,$$

then the distance between any two points of $\mathfrak{N}$ does not exceed $2(2rh + h^2)^{\frac{1}{2}}$.

Proof. If x_1 and x_2 are two points of $\mathfrak{N}$, the line segment between x_1 and x_2 certainly does not pierce the inner sphere. However, the entire line, of which this segment is a part, may or may not pierce this sphere. It is accordingly desirable to divide the proof into two cases depending upon whether the condition

$$(17) \qquad \| x_1 + k(x_2 - x_1) - \alpha \| \geqq r$$

is satisfied for all real values of k, or only for some including certainly $0 \leqq k \leqq 1$.

Case I. When (17) holds for all real k.

We denote by w the projection of the center α upon the line $\overline{x_1x_2}$ (Theorem 1), and by u the intersection of the projector $\overline{\alpha w}$ with the boundary of the sphere $\| x - \alpha \| \leqq r$. By Theorem 5 there exists through this boundary point u a supporting plane of the sphere, and its equation may be written

$$\pi(x) \equiv ((u - \alpha, x - u)) = 0.$$

Simple considerations then verify that

$$\pi(x_1) \geqq 0 \quad \text{and} \quad \pi(x_2) \geqq 0.$$

These results will be used after a consideration of Case II.

Case II. When (17) does not hold for all real k.

In this case there will be a parameter value $k = k_1$ for which (17) will not hold and a corresponding point

$$y = x_1 + k_1(x_2 - x_1)$$

of the line $\overline{x_1x_2}$ which lies in the inner sphere. The parameter value k_1 will be less than zero or greater than one according to the relative rôles played by x_1 and x_2. We will assume that $k_1 < 0$, and this means (geometrically speaking) that x_1 is closer than x_2 to the inner sphere.

We denote by u the intersection of the line $\overline{\alpha x_1}$ with the boundary of the inner sphere $\| x - \alpha \| \leqq r$, and write the equation of the supporting plane of the inner sphere through u in the form

$$\pi(x) \equiv ((u - \alpha, x - u)) = 0.$$

Again, simple considerations, depending upon the positions of α, y, x_1, and x_2 relative to the supporting plane and upon the assumed sign of k_1, enable us to show that

$$\pi(x_1) \geqq 0 \quad \text{and} \quad \pi(x_2) > 0.$$

In either Case I or Case II we find an upper bound for the distance $\| x_2 - x_1 \|$ as follows. Since

$$\| x_i - \alpha \|^2 = \| x_i - u \|^2 + \| u - \alpha \|^2 + 2((x_i - u, u - \alpha)) \qquad (i = 1, 2),$$

we have, by transposition and substitution,

$$\begin{aligned} \| x_i - u \|^2 &= \| x_i - \alpha \|^2 - r^2 - 2\pi(x_i) \\ &\leqq (r + h)^2 - r^2 = 2rh + h^2. \end{aligned}$$

Therefore we have, since

$$\| x_2 - x_1 \| \leqq \| x_2 - u \| + \| x_1 - u \|,$$

the desired conclusion

$$\| x_2 - x_1 \| \leqq 2(2rh + h^2)^{1/2}.$$

THEOREM 13. *If the linearly connected set $\mathfrak{M}$ is closed, every exterior point of $\mathfrak{M}$ has a unique projection on $\mathfrak{M}$.*

Proof. It is only necessary to show that every exterior point has a projection on $\mathfrak{M}$; it will then follow from Theorem 8 that this projection is unique.

Let y be an exterior point of $\mathfrak{M}$ whose distance from $\mathfrak{M}$ is equal to d; it is necessary to show the existence of a point x in $\mathfrak{M}$ for which $\| x - y \| = d$.

Select a strictly monotone decreasing sequence of positive numbers $\{\epsilon_n\}$ with limit zero. Corresponding to each ϵ_n there exists an x_n in $\mathfrak{M}$ such that

$$d \leqq \| x_n - y \| \leqq d + \epsilon_n .$$

Corresponding to any preassigned positive number ϵ, there exists an index m such that

$$2(2d\epsilon_m + \epsilon_m^2)^{\frac{1}{2}} < \epsilon,$$

and such that the points x_n for $n \geqq m$ all belong to the linearly connected aggregate common to $\mathfrak{M}$ and the sphere $\| x - y \| \leqq d + \epsilon_m$, which aggregate must in fact lie wholly within the spherical shell $d \leqq \| x - y \| \leqq d + \epsilon_m$.

Hence, by Theorem 12, we have

$$\| x_n - x_{n+p} \| \leqq 2(2d\epsilon_m + \epsilon_m^2)^{\frac{1}{2}} < \epsilon \qquad (n \geqq m; p = 1, 2, 3, \cdots).$$

The sequence $\{x_n\}$ therefore converges; and because of the completeness of the space $\mathfrak{S}$ and the closure of the set $\mathfrak{M}$, there exists an element x of $\mathfrak{M}$ such that $\lim_{n\to\infty} \| x - x_n \| = 0$. Now

$$x - y = (x - x_n) + (x_n - y),$$

and

$$\big| \| x - x_n \| - \| x_n - y \| \big| \leqq \| x - y \| \leqq \| x - x_n \| + \| x_n - y \| .$$

Since

$$\lim_{n\to\infty} \| x_n - y \| = d \qquad \text{and} \qquad \lim_{n\to\infty} \| x - x_n \| = 0,$$

it follows that

$$\| x - y \| = d.$$

In order to show that the linearly connected set $\mathfrak{M}$ is completely supported at its boundary, it would be necessary to show that each boundary point of $\mathfrak{M}$ is the projection of some exterior point. That this is not true even for linearly connected and closed sets can be illustrated by the following

Example. Let $\mathfrak{S}$ be the space of functions whose squares are Lebesgue integrable on the interval $0 \leqq t \leqq 1$. Let

$$((x, y)) \equiv \int_0^1 x(t)y(t)\, dt.$$

Let $\mathfrak{M}$ be the set for which $| x(t) | \leqq t$. It is easily verified that $\mathfrak{M}$ is linearly connected and closed. Also $x_0(t) \equiv 0$ is a boundary point of $\mathfrak{M}$, for it may be approached by the sequence of exterior points $\{y_n(t)\}$, where

$$y_n(t) = n^{-1} \qquad \text{for } 0 \leqq t \leqq n^{-1},$$

and

$$y_n(t) = 0 \qquad \text{for } n^{-1} < t \leqq 1.$$

But the set $\mathfrak{M}$ has no supporting plane through the boundary point $x_0(t)$. For suppose that there were such a supporting plane given by

$$\pi(x) \equiv \int_0^1 \alpha(t)x(t)\,dt = 0.$$

Since $\|\alpha\| \neq 0$, there is a set of points E of positive measure on the interval $0 \leqq t \leqq 1$ for which $\alpha(t) \geqq c > 0$ or else $\alpha(t) \leqq -c < 0$, where c is some fixed number. For definiteness assume the former case true. Consider the points $x_1(t)$ and $x_2(t)$ belonging to $\mathfrak{M}$ defined as follows:

$$x_1(t) = t \qquad \text{on } E \quad \text{and} \quad x_1(t) = 0 \qquad \text{on } C(E),$$

$$x_2(t) = -t \qquad \text{on } E \quad \text{and} \quad x_2(t) = 0 \qquad \text{on } C(E),$$

where $C(E)$ denotes the complement of E on the interval $0 \leqq t \leqq 1$. Then $\pi(x_1) > 0$ and $\pi(x_2) < 0$. Hence $\pi(x) = 0$ separates points of $\mathfrak{M}$. Therefore, there is no supporting plane through $x_0(t)$.

Though a linearly connected and closed set $\mathfrak{M}$ is not necessarily completely supported at its boundary, we can obtain a result on the distribution of the boundary points at which the set is supported. This result is given in the following theorem.

THEOREM 14. *The set of points at which a linearly connected and closed set $\mathfrak{M}$ is supported is everywhere dense on the boundary of $\mathfrak{M}$.*

Proof. It is necessary to show that if x_0 is any boundary point of $\mathfrak{M}$, then corresponding to any preassigned $\epsilon > 0$, there exists a normal point x_1 such that $\| x_1 - x_0 \| < \epsilon$. Let y_1 be an exterior point of $\mathfrak{M}$ such that $\| y_1 - x_0 \| < \epsilon$, and let the projection of y_1 on $\mathfrak{M}$ be x_1 (Theorem 13); then x_1 is a *normal* point (Theorem 11). The plane $\pi(x) = 0$, where

$$\pi(x) \equiv ((y_1 - x_1, x - x_1)),$$

is a supporting plane of $\mathfrak{M}$ through x_1. Since $\pi(y_1) > 0$, we must have $\pi(x_0) \leqq 0$. Now

$$y_1 - x_0 = (y_1 - x_1) - (x_0 - x_1),$$

and

$$\| y_1 - x_0 \|^2 = \| y_1 - x_1 \|^2 - 2\pi(x_0) + \| x_0 - x_1 \|^2.$$

Hence,

$$\| x_0 - x_1 \|^2 = \| y_1 - x_0 \|^2 - \| y_1 - x_1 \|^2 + 2\pi(x_0),$$

from which it follows that

$$\| x_0 - x_1 \|^2 < \| y_1 - x_0 \|^2 < \epsilon^2;$$

consequently,

$$\| x_0 - x_1 \| < \epsilon.$$

THEOREM 15. *Let x_0 be any boundary point of a linearly connected and closed set $\mathfrak{M}$, and let $\{y_n\}$ be a sequence of exterior points having x_0 as their limit point. Let x_n be the projection of y_n on $\mathfrak{M}$, and let*

$$\alpha_n = \frac{y_n - x_n}{\| y_n - x_n \|}. \tag{18}$$

Then if[5] *from the sequence $\{\alpha_n\}$ a subsequence $\{\alpha_{n_k}\}$ can be selected so that for some element α in $\mathfrak{S}$ we have*

$$\lim_{k \to \infty} ((\alpha_{n_k}, x)) = ((\alpha, x)) \qquad \textit{for every } x \textit{ in } \mathfrak{S} \tag{19}$$

and

$$\| \alpha \| \neq 0, \tag{20}$$

then x_0 is a normal point of $\mathfrak{M}$.

Proof. The planes

$$\pi_k(x) \equiv ((\alpha_{n_k}, x - x_{n_k})) = 0$$

are supporting planes of $\mathfrak{M}$ (Theorem 10). Let y be any point of $\mathfrak{M}$; then

$$\pi_k(y) \leqq 0 \quad \text{for each } k. \tag{21}$$

We may write

$$\pi_k(y) = ((\alpha_{n_k}, y - x_0)) + \pi_k(x_0). \tag{22}$$

Since

$$\lim_{k \to \infty} ((\alpha_{n_k}, y - x_0)) = ((\alpha, y - x_0))$$

by (19) and

$$\lim_{k \to \infty} \pi_k(x_0) = 0,$$

we have from (22)

$$\lim_{k \to \infty} \pi_k(y) = ((\alpha, y - x_0)).$$

But because of (21), it follows that

$$((\alpha, y - x_0)) \leqq 0.$$

Consequently,

$$\pi_0(x) \equiv ((\alpha, x - x_0)) = 0$$

is a supporting plane of $\mathfrak{M}$ through x_0, since $\pi_0(y) \leqq 0$ for every y in $\mathfrak{M}$. Therefore, x_0 is a normal point of $\mathfrak{M}$.

[5] In particular this condition will be satisfied if the aggregate of boundary points of the unit sphere in $\mathfrak{S}$ is *weakly compact* (Banach, *Théorie des Opérations Linéaires*, p. 239).

THEOREM 16. *Let x_0 be any boundary point of a linearly connected and closed set $\mathfrak{M}$, and let $\{y_n\}$ be a sequence of exterior points having x_0 as their limit point. Let x_n be the projection of y_n on $\mathfrak{M}$ and let α_n be given by* (18). *Then, if from the sequence $\{\alpha_n\}$ a subsequence $\{\alpha_{n_k}\}$ can be selected so that for some α in $\mathfrak{S}$ we have*

$$\lim_{k\to\infty} \| \alpha_{n_k} - \alpha \| = 0, \tag{23}$$

x_0 is a normal point of $\mathfrak{M}$.

Proof. The condition (23) implies (19) and (20). Hence the hypotheses of Theorem 15 are satisfied, and x_0 is a normal point of $\mathfrak{M}$.

From the preceding theorem, the following corollaries are immediately deducible.

COROLLARY 1. *A linearly connected and closed set $\mathfrak{M}$ in a compact space is completely supported at its boundary.*

COROLLARY 2. *A linearly connected and closed set $\mathfrak{M}$ in a finite-dimensioned space is completely supported at its boundary.*

CARNEGIE INSTITUTE OF TECHNOLOGY.

FIXED SETS UNDER HOMEOMORPHISMS

By J. L. Kelley

As a generalization of the Scherrer Fixed-Point Theorem[1] W. L. Ayres has shown[2] that a homeomorphism T which carries a locally connected continuum M into a subset of itself carries some cyclic element C of M into a subset of itself. This result yields a fixed point theorem in an acyclic space or under certain other conditions. In this paper the following extension of Ayres' theorem to non-locally connected continua will be proved:

Theorem I. *If T is a homeomorphism carrying a compact continuum M into a subset of itself, then there exists a subcontinuum H of M, with $T(H) = H$ and such that H contains no cut points of itself.*

This will be deduced as a consequence of the following series of lemmas.

Lemma I. *If M is a compact continuum and T is a continuous transformation, $T(M) \subset M$, then $\pi = \prod_{i=1}^{\infty} T^i(M)$ is an invariant continuum under T, i.e., $T(\pi) = \pi$.*

Proof. Since $T(M) \subset M$, $T^{i+1}(M) \subset T^i(M)$ and therefore π is a continuum. If $p \in \pi$, $T^{-1}(p)$ intersects every set $T^i(M)$ and hence also π. Therefore $p \in T(\pi)$ and $\pi \subset T(\pi)$. If $p \in \pi$, $p \in T^i(M)$ for all $i > 0$ and hence $T(p) \in T^{i+1}(M)$. Therefore $T(p) \in \pi$ and $T(\pi) \subset \pi$.

Without further proof we may also state

Lemma I'.[3] *If T is a homeomorphism and $T^{-1}(M) \subset M$, other conditions being the same as in Lemma I, then $\pi = \prod_{i=1}^{\infty} T^{-i}(M)$ is an invariant continuum under T.*

Lemma II. *The property of being an invariant continuum under a continuous transformation T is inducible.*

Proof. Suppose $\pi = \prod_{i=1}^{\infty} C_i$ with C_i a continuum, $T(C_i) = C_i$ and $C_{i+1} \subset C_i$ for all $i > 0$. Then "p a point of π" implies "p, and hence $T(p)$, in all C_i"; thus $T(p) \in \pi$. It follows that $T(\pi) \subset \pi$. Furthermore, the inverse of every point $p \in \pi$ intersects every C_i and hence intersects π. Therefore $p \in T(\pi)$ and $\pi \subset T(\pi)$.

Received January 20, 1939.

[1] *Über Ungeschlossene stetige Kurven*, Mathematische Zeitschrift, vol. 24(1925), pp. 125–130.

[2] *Some generalizations of the Scherrer Fixed Point Theorem*, Fundamenta Mathematicae, vol. 16(1930), pp. 332–336. See also K. Borsuk, *Einige Sätze über stetige Streckenbilder*, Fundamenta Mathematicae, vol. 18(1932), pp. 198–213.

[3] If T is not a homeomorphism, the process defined here leads, by a lemma similar to Lemma II, to a collection of invariant sets not necessarily connected. These will be subsets of (not necessarily irreducibly) invariant continua.

As a result of the previous lemmas we have, by the Brouwer Reduction Theorem, the following.

LEMMA III.[4] *If M is a continuum and T is a continuous transformation $T(M) \subset M$, then there exists a continuum $N \subset M$ invariant under T and irreducible with respect to the property of being an invariant continuum under T.*

We remark that, from Lemma I, it is impossible that the continuum N of Lemma III contain a proper subcontinuum C such that $T(C) \subset C$, and if T is a homeomorphism, it is also impossible that $T(C) \supset C$.

LEMMA IV. *If $T(M) \subset M$ is a homeomorphism, no irreducibly invariant subcontinuum of M can contain a cut point of itself.*

Proof. Suppose the contrary. Let N be an irreducibly invariant subcontinuum of M with $N = A + B$, and A and B continua, the product AB being a single point p. Suppose $T(p) \epsilon A$. For $x \epsilon A$, define $\phi(x) = T(x)$ provided $T(x) \epsilon A$, and $\phi(x) = p$ if $T(x) \epsilon B$. Then ϕ is a continuous transformation and $\phi(A) \subset A$. Let $\pi = \prod_{i=1}^{\infty} \phi^i(A)$. Then π is an invariant continuum under ϕ. Suppose first that $\pi \not\supset p$. Then on π, ϕ is identical with T and π is fixed under T, and this is a contradiction. Suppose then that $p \epsilon \pi$; then also $T(p) \epsilon \pi$. $T(\pi)$ is a continuum containing $T(p) \epsilon A$, and since $T(\pi)$ contains all points of $\pi - p$, it contains also p. Hence $T(\pi) \supset \pi$. But it is impossible that any subcontinuum of N be contained in its transform. We have then a contradiction.

The theorem then follows without further proof.

We now make some applications of our theorem. According to G. T. Whyburn[5] if M is a continuum, a subcontinuum N of M will be said to be a *0-th order cyclic element* of M, or simply an *E_0 set* provided N is a maximal with respect to the property of being a subcontinuum without cut points. We may now state

THEOREM II. *If M is a continuum and T is a homeomorphism, $T(M) \subset M$, then there exists either a fixed point in M or else a set $E_0 \subset M$ such that $T(E_0) \subset E_0$.*

Proof. An irreducibly invariant subcontinuum H of M is evidently either a point or is contained[5] in a set E_0. If H is non-degenerate, then, since the transform of the set E_0 containing H is a set of the same type, it follows that $T(E_0) \subset E_0$.

We shall say that a set H has the *fixed point property for homeomorphisms* provided every homeomorphism $T(H) \subset H$ leaves one point fixed. From the previous theorem, a homeomorphism on a continuum M, $T(M) \subset M$, implies the existence of a set E_0 such that $T(E_0) \subset E_0$. We may then state

[4] It is to be noted that Lemmas I, II and III are proved for any continuous transformation.

[5] See *Concerning maximal sets*, Bulletin of the American Mathematical Society, vol. 40(1934), pp. 159–164; also *Cyclic elements of higher order*, American Journal of Mathematics, vol. 56(1934), pp. 133–146.

THEOREM III.[6] *If every set E_0 in a continuum M has the fixed point property for homeomorphisms, so also has M; i.e., this fixed point property is E_0 extensible.*

Since, if M is locally connected, the sets E_0 are simply the cyclic elements of M, we may deduce from the two previous theorems the following known theorems (see footnote 1).

THEOREM IV. *If a homeomorphism T carries a locally connected continuum M into a subset of itself, then there exists a cyclic element C of M such that $T(C) \subset C$.*

THEOREM V. *The fixed point property for homeomorphisms is cyclically extensible.*

UNIVERSITY OF VIRGINIA.

[6] This result appears to be related to that of O. H. Hamilton, who has shown that a hereditarily unicoherent and hereditarily decomposable continuum has the fixed point property for homeomorphisms (see Transactions of the American Mathematical Society, vol. 44(1938), pp. 18-24). However, Hamilton's theorem neither includes nor is included in Theorem III, for one can construct a hereditarily unicoherent and hereditarily decomposable continuum which is without cut points. Hamilton's theorem indicates the fixed point property for homeomorphisms for this continuum, while Theorem III gives no information. On the other hand, Hamilton's theorem applies necessarily to 1-dimensional continua, while here there is no such restriction.

THE EXTENSION OF LINEAR FUNCTIONALS

By A. E. Taylor

1. **Introduction.** If E is a Banach space, its conjugate space E^*, the space of all linear functionals defined on E, is also a Banach space. If $\mathfrak{M}$ is a linear manifold in E, and φ is a linear functional on $\mathfrak{M}$, regarded as a normed linear space in itself, a well known theorem[1] asserts that it is possible to extend φ to all of E without increasing its norm. The extension thus obtained will not, in general, be unique. One of the considerations of this paper is the establishment of criteria for the uniqueness of such extensions in all cases—i.e., for *all* subspaces $\mathfrak{M}$ of the given space. A sufficient condition is found to be that the unit sphere in E^* be strictly convex; this condition is also necessary in case E is reflexive (for details see §4).

Another question which arises is that of whether a rule of extension may be established which will be linear. That is, is it possible to define a *linear* operation A on $\mathfrak{M}^*$ to E^* such that the linear functional $f = A\varphi$ will be an extension of φ, for each φ in $\mathfrak{M}^*$? This will imply, as we show in §3, that $\mathfrak{M}^*$ is isomorphic[2] with a linear subspace in E^*. This question is for the most part distinct from the question of uniqueness of extension and is discussed in §3 with the aid of the notion of a projection of a space on a subspace. §5 is devoted to a few remarks on a situation in which the uniqueness of extension and the existence of the operation A, described above, are bound together. Finally, in §6 an example is given to show the irredundancy of a part of the hypothesis of Theorem 3.

2. **Notation.** Throughout the paper the following conventions in notation will be observed. E denotes a Banach space, E^* its conjugate space, and E^{**} the space conjugate to E^*. $\mathfrak{M}$ denotes a closed linear manifold in E, $\mathfrak{M}^*$ the space conjugate to $\mathfrak{M}$ considered as a space by itself. We use letters $x, y, \cdots$ for elements of E; $f, g, \cdots$ for elements of E^*; $F, G, \cdots$ for elements of E^{**}. Elements of $\mathfrak{M}^*$ are denoted by φ, and elements of $\mathfrak{M}^{**}$ by Φ.

For our purposes it is frequently convenient to introduce the following notation for the values of various linear functionals. We write

$$f(x) = (x, f),$$
$$\varphi(x) = [x, \varphi],$$
$$F(f) = \{f, F\},$$
$$\Phi(\varphi) = \langle \varphi, \Phi \rangle,$$

Received January 26, 1939.

[1] S. Banach, *Opérations Linéaires*, p. 55, Théorème 2. We shall refer to this as the Hahn-Banach theorem.

[2] For the definition of this term see the beginning of §3.

where the letters have the significance explained above. It should be kept in mind that (x, f) is a bilinear functional of the arguments x, f, and that $\|f\|$ is the smallest constant C such that $|(x, f)| \leqq C \|x\|$ for all $x \in E$. Similar remarks apply to the other expressions.

We denote by T the linear operation on E^* to $\mathfrak{M}^*$ defined by the condition

$$(2.1) \qquad [x, Tf] = (x, f), \qquad x \in \mathfrak{M}, f \in E^*.$$

Evidently $\|Tf\| \leqq \|f\|$. But, by the Hahn-Banach theorem (see footnote 1), if $\varphi \in \mathfrak{M}^*$, there exists an $f \in E^*$ such that $Tf = \varphi$ and $\|f\| = \|\varphi\|$. Hence $\|T\| = 1$ and the range $T(E^*)$ of T is precisely $\mathfrak{M}^*$.

3. **The relationship of $\mathfrak{M}^*$ and E^*.** We first recall the notions of isomorphism and equivalence for normed linear spaces. E_1 and E_2 are said to be *isomorphic* if there exists an operation A which carries E_1 into all of E_2 in a one-to-one manner, and if further A and its inverse A^{-1} are linear operations.[3] The spaces are said to be equivalent if in addition $\|Ax\| = \|x\|$ for each x in E_1.

In discussing the relationship of $\mathfrak{M}^*$ and E^* we shall assume that $\mathfrak{M}$ is a *closed* linear manifold. There is no loss of generality in so doing, for if $\mathfrak{M}$ were not closed, we could consider its closure $\overline{\mathfrak{M}}$, and the conjugate space $(\overline{\mathfrak{M}})^*$ would be equivalent to $\mathfrak{M}^*$, as is easily seen.

Finally, we shall define the notion of a projection of a space E. By a projection of E is meant a linear operation P on E to E with the property $P^2 = P$. It is easily seen that the range of P is characterized by the fact that $Px = x$ for elements of this range. Hence the range is a closed linear manifold. If the range is $\mathfrak{M}$, we say that P projects E on $\mathfrak{M}$. The conjugate operation P^* on E^* to E^*, defined by the condition $(x, P^*f) = (Px, f)$, $x \in E, f \in E^*$ is a projection of E^*, as may be verified at once. Its range is the class of f such that $(x, f) = (Px, f)$ for each x in E.

THEOREM 1. *If P is a projection of E on $\mathfrak{M}$, there exists a linear operation A on $\mathfrak{M}^*$ to E^* with the following properties:*

(1) *the range of A is the range $P^*(E^*)$ of P^*, and A sets up an isomorphism between $\mathfrak{M}^*$ and $P^*(E^*)$;*

(2) $[x, \varphi] = (x, A\varphi)$ *if* $x \in \mathfrak{M}, \varphi \in \mathfrak{M}^*$;

(3) $P^* = AT$; $\|A\| = \|P^*\| = \|P\|$.

Proof. Define A by the condition $(x, A\varphi) = [Px, \varphi]$, which condition is possible since $Px \in \mathfrak{M}$. Clearly A is linear, with $\|A\| \leqq \|P\|$. Since $Px = x$ in $\mathfrak{M}$, property (2) is evident. AT is an operation on E^* to E^* defined by $(x, ATf) = [Px, Tf] = (Px, f) = (x, P^*f)$. Therefore $AT = P^*$. Since the range of T is exactly $\mathfrak{M}^*$, we infer that the ranges of A and P^* are identical. From (2) it follows that if $A\varphi = 0$, then $\varphi = 0$. This means that A describes a one-to-one correspondence between $\mathfrak{M}^*$ and $P^*(E^*)$. Since both of these

[3] By a linear operation we shall mean one which is distributive and continuous.

sets are *complete* linear spaces, the inverse map is also continuous,[4] and the mapping describes an isomorphism. Finally, from $P^* = AT$ we have $\|P^*\| \leqq \|A\| \cdot \|T\| = \|A\|$. Because of the relation[5] $\|P^*\| = \|P\|$ we conclude the last part of (3).

Corollary. *If $\|P\| = 1$, A sets up an equivalence between $\mathfrak{M}^*$ and $P^*(E^*)$.*

For then certainly $\|A\varphi\| \leqq \|\varphi\|$. But $TA\varphi = \varphi$ for all $\varphi \in \mathfrak{M}^*$ because $[x, TA\varphi] = (x, A\varphi) = [x, \varphi]$ if $x \in \mathfrak{M}$. Therefore $\|\varphi\| \leqq \|A\varphi\|$, and so $\|\varphi\| = \|A\varphi\|$.

Theorem 2. *If there exists a linear operation A on $\mathfrak{M}^*$ to E^* with the property*

$$[x, \varphi] = (x, A\varphi), \qquad x \in \mathfrak{M}, \varphi \in \mathfrak{M}^*, \tag{3.1}$$

the range $A(\mathfrak{M}^*)$ of A is a closed linear manifold in E^*, and there exists a projection Q of E^* on $A(\mathfrak{M}^*)$. Furthermore, A sets up an isomorphism between $\mathfrak{M}^*$ and the space $A(\mathfrak{M}^*)$, and if A^{-1} denotes the inverse operation, then

$$\frac{1}{\|A^{-1}\|} \leqq \|Q\| \leqq \|A\|. \tag{3.2}$$

Proof. We define $Q = AT$. Then Q is linear, and $\|Q\| \leqq \|A\|$. It is also immediate that the ranges of Q and A are identical, for the range of T is exactly $\mathfrak{M}^*$. That Q is a projection follows from the fact that $TA\varphi = \varphi$ for all $\varphi \in \mathfrak{M}^*$; this is because $[x, TA\varphi] = (x, A\varphi) = [x, \varphi]$ by the properties of A and T. Thus the common range of A and Q is a closed linear manifold in E^*. Since $TA\varphi = \varphi$, we have $\|\varphi\| \leqq \|A\varphi\|$, and so $A\varphi = 0$ implies $\varphi = 0$. As in the proof of the previous theorem this implies that A defines an isomorphism of $\mathfrak{M}^*$ and $A(\mathfrak{M}^*)$. The relation $A^{-1}Q = T$ then leads to the inequality $1 = \|T\| \leqq \|A^{-1}\| \cdot \|Q\|$. This completes the proof.

Corollary. *If A defines an equivalence of $\mathfrak{M}^*$ and $A(\mathfrak{M}^*)$, $\|A\| = \|A^{-1}\| = 1$, and so $\|Q\| = 1$.*

Since a projection P of E automatically induces a projection P^* of E^*, the question naturally arises as to whether the operation Q of Theorem 2, which is a projection of E^*, may not induce a projection of E on $\mathfrak{M}$. The natural way to investigate this matter is to observe that Q^* is a projection of E^{**}, and to proceed from this, using the well known fact that E may be imbedded in E^{**}. This method has been carried through in Theorem 3, below, with the aid of the additional assumption that $\mathfrak{M}$ is reflexive. That this assumption is essential in order to obtain the stated conclusions is shown by an example in §6.

Definition. A Banach space E is said to be reflexive if every element F of E^{**} is representable in the form $\{f, F\} = (x, f)$ for some x in E and all f in E^*.

[4] Banach, op. cit., p. 41, Théorème 5.

[5] Banach, op. cit., p. 100, Théorème 3.

When E is reflexive, the correspondence $F \leftrightarrow x$ defines an equivalence[6] of E and E^{**}. It is further known that if E is reflexive, the same is true of all its closed linear subspaces.[7] However, a non-reflexive space can have closed subspaces which are reflexive. For example, any finite dimensional subspace is reflexive.

THEOREM 3. *Under the assumptions of Theorem 2 and the further assumption that $\mathfrak{M}$ is reflexive there exists a projection P of E on $\mathfrak{M}$, with the properties:*

(1) $P^* = Q$;

(2) *the operation A of Theorem 2 is given by $(x, A\varphi) = [Px, \varphi]$, and $\| A \| = \| P \| = \| Q \|$.*

Proof. If we apply Theorem 1 to the space E^* and the projection Q of this space, we infer, because of the fact that $\mathfrak{M}^*$ is isomorphic[8] with the range of Q, that $\mathfrak{M}^{**}$ is isomorphic with the range of the projection Q^* of E^{**}. Therefore, if we think of E as being imbedded in E^{**} and assume that $\mathfrak{M}$ is reflexive, so that $\mathfrak{M}$ and $\mathfrak{M}^{**}$ are equivalent, it is easy to see how Q^* defines an operation on E to $\mathfrak{M}$. With this outline of the procedure in mind it is merely a matter of setting up the necessary notational machinery to carry out the details.

First we define the imbedding of E in E^{**} by the operation V on E to E^{**} given by

$$\{f, Vx\} = (x, f), \qquad x \in E, f \in E^*.$$

This defines a linear isometric (see footnote 6) ($\| Vx \| = \| x \|$) imbedding of E in E^{**}. V defines an equivalence of E and E^{**} if and only if E is reflexive.

Next, $Q = AT$, and so $Q^* = T^*A^*$, by the rules for forming conjugate operations. Now A^* is an operation on E^{**} to $\mathfrak{M}^{**}$, given by

$$\langle \varphi, A^*F \rangle = \{A\varphi, F\}.$$

The range of A^* is exactly $\mathfrak{M}^{**}$, for any functional Φ defined on $\mathfrak{M}^*$ generates, by the operation A, a linear functional on the linear subset $A(\mathfrak{M}^*)$ in E^*, and this latter functional has an extension to all of E^*, by the Hahn-Banach theorem (see footnote 1). Thus A^* plays the same rôle with respect to E^*, $A^*(\mathfrak{M}^*)$ as T plays with respect to E, $\mathfrak{M}$. The equation $Q^* = T^*A^*$ now shows that the ranges of Q^* and T^* are identical. But more is true. T^* sets up a one-to-one correspondence between $\mathfrak{M}^{**}$ and the range of Q^*, and therefore $\mathfrak{M}^{**}$ and $Q^*(E^{**})$ are isomorphic. To see this we have only to prove that $T^*\Phi = 0$ implies $\Phi = 0$. Now

$$\{f, T^*\Phi\} = \langle Tf, \Phi \rangle,$$

[6] For it follows from Banach, op. cit., p. 55, Théorème 3, that $\sup |f(x)| = \| x \|$ (for all $\| f \| = 1$), so that $\| F \| = \| x \|$. Even if E is not reflexive, this defines a linear imbedding of E in E^{**}.

[7] B. J. Pettis, Bulletin of the American Mathematical Society, vol. 44(1938), p. 423, Theorem 3.

[8] The conjugate spaces of isomorphic spaces are isomorphic: see Banach, op. cit., p. 188.

and as f runs over E^*, Tf runs over all of $\mathfrak{M}^*$, and the above assertion is verified. Note that $A^*T^*\Phi = \Phi$ for all Φ in $\mathfrak{M}^{**}$. For

$$\langle \varphi, A^*T^*\Phi \rangle = \{A\varphi, T^*\Phi\} = \langle TA\varphi, \Phi \rangle,$$

and we saw in the proof of Theorem 2 that $TA\varphi = \varphi$.

Finally, if $\mathfrak{M}$ is reflexive, there is a unique operation U on $\mathfrak{M}^{**}$ to $\mathfrak{M}$ such that $\langle \varphi, \Phi \rangle = [U\Phi, \varphi]$. We now define the operation P on E to $\mathfrak{M}$ by the equation $P = UA^*V$. It is evidently linear. It is a projection, for

$$\{f, VU\Phi\} = (U\Phi, f) = [U\Phi, Tf] = \langle Tf, \Phi \rangle = \{f, T^*\Phi\},$$

that is, $VU = T^*$, and therefore $P^2 = UA^*T^*A^*V = UA^*V = P$, since A^*T^* is the identity transformation of $\mathfrak{M}^{**}$. Now, if x is in E, $[Px, \varphi] = [UA^*Vx, \varphi] = \langle \varphi, A^*Vx \rangle = \{A\varphi, Vx\} = (x, A\varphi)$. But if x is in $\mathfrak{M}$, $(x, A\varphi) = [x, \varphi]$, and therefore $Px = x$. The range of P is therefore exactly $\mathfrak{M}$, and we conclude from Theorem 2 that $\| P \| = \| A \|$. But also,

$$(x, P^*f) = (Px, f) = [Px, Tf] = (x, ATf) = (x, Qf)$$

by the above reckoning. Thus $P^* = Q$ and the proof is complete.

Remark. The assumptions of Theorem 2 are not fulfilled for all the closed linear manifolds in all Banach spaces, or even in all reflexive spaces. For it can happen that $\mathfrak{M}$ is such that there is no projection of E on $\mathfrak{M}$ (see §6).

As an instance of the situation set forth in Theorem 2 we note that if we regard E as being imbedded in E^{**}, and E^* as imbedded (in the same manner) in E^{***}, then there is a projection, of unit norm, of E^{***} on the map of E^* in E^{***}. Looked at in another way, this amounts to saying that the identity operation of E^* on itself admits a linear extension, as an operation on E^{***} to E^*, without increase of norm. That the like statement is in general not true of E^{**} and E is shown by the example in §6. There it is shown that if a projection of (m) on (c_0) exists, it is of norm $\geqq 2$. Since (m) is equivalent to $(c_0)^{**}$, this furnishes us with a counter example in the present situation.

4. **Unique extension of linear functionals.** In this section the assumption that E is a complete space is unnecessary.

Definition. The space E is said to have property A if, $\mathfrak{M}$ being an arbitrary closed linear manifold in E, not E itself, and φ an arbitrary element of $\mathfrak{M}^*$, there exists a unique element f of E^* such that $Tf = \varphi$, $\| f \| = \| \varphi \|$, where T is the operation defined in §2. (It suffices to consider $\mathfrak{M} \neq E$, for if $\mathfrak{M} = E$ the above requirement is automatically satisfied.)

A closely related property of a space E is described in the following definition.

Definition. The space E is said to have property B if for each element $x_0 \neq 0$ in E there is a unique element f of E^* such that $\| f \| = 1, f(x_0) = \| x_0 \|$.

The Hahn-Banach theorem (see footnote 1) guarantees the existence of *at*

least one f to meet the conditions laid down in the definitions. The uniqueness is the essential thing.

It is clear that if E has property A, it also has property B.

THEOREM 4. *If E has property* B *and if every closed linear manifold $\mathfrak{M}$, $\mathfrak{M} \neq E$, in E is reflexive (as, for example, when E itself is reflexive), then E also has property* A.

Proof. First note that if $\mathfrak{M}$ is reflexive and φ is in $\mathfrak{M}^*$, then, for some x_0 in $\mathfrak{M}$, $\| x_0 \| = 1$ and $\varphi(x_0) = \| x_0 \|$. For there exists[9] $\Phi \in \mathfrak{M}^{**}$, $\| \Phi \| = 1$ and $\Phi(\varphi) = \| \varphi \|$. If $\mathfrak{M}$ is reflexive, however, we can write $\Phi(\varphi)$ in the form $\varphi(x)$ with $x \in \mathfrak{M}$, $\| x \| = \| \varphi \|$. Now suppose that $f_1, f_2 \in E^*$, with $f_1 \neq f_2$, $\| f_1 \| = \| f_2 \| = \| \varphi \|$ and $Tf_1 = Tf_2 = \varphi$. Define $g_1 = (1/\| \varphi \|)f_1$, $g_2 = (1/\| \varphi \|)f_2$. Then $\| g_1 \| = \| g_2 \| = 1$, and $g_1(x_0) = g_2(x_0) = 1 = \| x_0 \|$, but $g_1 \neq g_2$, contrary to property B.

Property B may be interpreted geometrically as follows: a set of elements S in E is called a hyperplane if and only if it consists of the totality of elements x satisfying an equation of the form $f(x) = c$. Two points *not* on a hyperplane S are said to be on the same side of it if the line segment joining them does not meet S. A hyperplane S is said to support the unit sphere $\| x \| = 1$ if and only if the distance from S to the unit sphere is zero, and all points for which $\| x \| \leqq 1$, x not on S, are on the same side of S. It may then be shown that a hyperplane supports the unit sphere if and only if it is given by $f(x) = 1$, where $\| f \| = 1$. Through each point of the unit sphere there passes at least one supporting hyperplane. If there is at most one, we shall call it a *tangent* hyperplane. Property B is then the property that the unit sphere of E admits, at each point, a tangent hyperplane.[10]

Another geometrical property of importance in the present connection is that of the strict convexity of the unit sphere.

DEFINITION. The unit sphere in E is said to be *strictly* convex provided that $\| x \| = 1$, $\| y \| = 1$, $x \neq y$, $0 < t < 1$ imply that $\| tx + (1 - t)y \| < 1$.

This means that the line segment joining two points on the surface of the sphere lies, with the exception of its end points, entirely inside the sphere. Every sphere in E will then be strictly convex also.

THEOREM 5. *If E^* has property* B, *the unit sphere in E is strictly convex.*

The proof depends on the following lemma.

[9] See the reference in footnote 6. Instead of assuming that every $\mathfrak{M}$ is reflexive it suffices to assume that the unit sphere $\| x \| = 1$ in each subspace $\mathfrak{M}$ is weakly compact. For if $\| x_n \| = 1$ and $\varphi(x_n) \to \| \varphi \|$, we can then find $x_0 \in \mathfrak{M}$ such that $\varphi(x_n) \to \varphi(x_0) = \| \varphi \|$. It follows that $\| x_0 \| \geqq 1$; since a convex, closed set is weakly closed, we conclude $\| x_0 \| = 1$. The proof is then as before.

[10] A general discussion of convex hypersurfaces and their supporting hyperplanes in normed linear spaces is given by S. Mazur, Studia Mathematica, vol. 4(1933), pp. 70–84.

LEMMA. *If the unit sphere in E is not strictly convex, there exists a line segment* $tx + (1 - t)y$, $0 \leq t \leq 1$, *all of whose points lie on the surface of the sphere; i.e.,* $\| tx + (1 - t)y \| = 1$.

Proof. There exist distinct collinear points x_1, x_2, x_3, with $\| x_i \| = 1$. Suppose that x_2 lies between x_1 and x_3. Unless all points of the segment (x_1, x_3) are of norm one (which is what has to be proved), there exist y between x_1 and x_2 and z between x_2 and x_3, with either $\| y \| \leq 1$, $\| z \| < 1$, or $\| y \| < 1$, $\| z \| \leq 1$. In either case $x_2 = ty + (1 - t)z$, $0 < t < 1$, $\| x_2 \| \leq t \| y \| + (1 - t) \| z \| < 1$, a contradiction.[11]

We now return to Theorem 5. If it is false, we may suppose that x_1, x_2 and the line segment joining them lie on the surface of the unit sphere. That is, $\| tx_1 + (1 - t)x_2 \| = 1$, $0 \leq t \leq 1$. Then $\| tx_1 + (1 - t)x_2 \| \geq 1$ for *all* real t. We can choose $f_0 \in E^*$ so that $\| f_0 \| = 1$, $f_0(x_i) = 1$ $(i = 1, 2)$. For the condition that such an f_0 exist is that[12]

$$| a + b | \leq \| ax_1 + bx_2 \|$$

for all numbers a, b. If we let $a + b = r$, the above is trivially true when $r = 0$. If $r \neq 0$, $\| ax_1 + bx_2 \| = \| ax_1 + (r - a)x_2 \| = | r | \, \| (a/r)x_1 + (1 - a/r)x_2 \| \geq | r | = | a + b |$, by the properties of x_1, x_2. If now we define $F_i \in E^{**}$ by $F_i(f) = f(x_i)$, we have $\| F_i \| = 1$ and $F_i(f_0) = 1 = \| f_0 \|$. But $F_1 \neq F_2$ if $x_1 \neq x_2$, and so E^* cannot have property B.

THEOREM 6. *If the unit sphere in* E^* *is strictly convex, E has property* A. *If E is reflexive, the converse is also true.*

Proof. Let $\mathfrak{M} \neq E$, φ be given, and suppose that $f, g \in E^*$, $\| f \| = \| g \| = \| \varphi \|$, $Tf = Tg = \varphi$. Then $T(tf + (1 - t)g) = t\varphi + (1 - t)\varphi = \varphi$ and hence $\| \varphi \| \leq \| tf + (1 - t)g \|$. But if $0 < t < 1$ and $f \neq g$, $\| tf + (1 - t)g \| < \varphi$, since the sphere of radius $\| \varphi \|$ is also strictly convex. Hence we must conclude $f = g$.

This is the direct part of the theorem. If E is reflexive, and if it has property A, then E^{**} also has property A (since E and E^{**} are equivalent). E^{**} therefore has property B, and this implies, by Theorem 5, that the unit sphere in E^* is strictly convex.

THEOREM 7. *If the unit sphere in E is strictly convex, and if E is reflexive,* E^* *has property* A (*which is the same here as property* B, *since* E^* *is also reflexive*).

Proof. Because of reflexivity it follows that the unit sphere in E^{**} is strictly convex. The conclusion then follows from Theorem 6.

It is natural to ask the following questions. Does property A, or the strict convexity of the unit sphere in E, imply the corresponding property for E^*? Are these properties dependent upon each other; that is, does either of them, possibly in conjunction with reflexivity, imply the other for the same space E?

[11] This proof of the lemma was suggested by the referee, as was also Theorem 7, below.

[12] Banach, op. cit., p. 57, Théorème 5.

The answers are in the negative, as may be shown by examples. To illustrate we consider a two-dimensional space of elements $x = (x_1, x_2)$, where the numbers x_1, x_2 are the Cartesian coördinates of the end point of the vector x. The conjugate space E^* is also two-dimensional, and its elements may be represented on the same coördinate system that is used for E. The "unit sphere" in each case is defined by a convex curve symmetric about the origin. By an examination of the lines of support for the curve for E one can determine, using the geometrical interpretation of property B, the nature of the curve for E^*. In doing this it is found that a *corner* on the curve for E (i.e., a place where the supporting line is not unique) leads to a flat portion of the curve for E^*. An example in which the curve for E is strictly convex, but with corners, while the curve for E^* is *not* strictly convex, and is without corners, is the following. We give the definitions of those parts of the curves in the first quadrant only.

E: $$(x_1 + \tfrac{1}{2})^2 + (x_2 + \tfrac{1}{2})^2 = \tfrac{5}{2},$$

$$x_2 = 1, \qquad 0 \leqq x_1 \leqq \tfrac{1}{3}.$$

E^*: $$9(x_1^2 + x_2^2) - 2x_1x_2 - 4(x_1 + x_2) - 4 = 0 \begin{pmatrix} \tfrac{1}{3} \leqq x_1 \leqq 1 \\ \tfrac{1}{3} \leqq x_2 \leqq 1 \end{pmatrix},$$

$$x_1 = 1, \qquad 0 \leqq x_2 \leqq \tfrac{1}{3}.$$

Thus E is reflexive (because finite dimensional), its unit sphere is strictly convex, but it does not have property A; E^* is reflexive, has property A, but its unit sphere is not strictly convex. Since $E^{**} = E$, this example shows that the answers to the above questions are in the negative.

We shall conclude this section with a few remarks about the converse part of Theorem 6. In the first place, note that if E has property A, so do all of its linear subspaces. Now suppose that the unit sphere in E^* is not strictly convex. Then evidently there is a two-dimensional linear subspace S of E^* whose unit sphere is likewise not strictly convex. If S were the conjugate space of some linear subspace in E, this subspace could not have property A, for S is reflexive. Therefore E itself could not have property A. It seems doubtful, however, that S need be the conjugate of any subspace of E.

5. **Further questions.** In this section we shall consider briefly the following question: Is there any connection between the property A for a space E and the property that for *every* closed linear manifold $\mathfrak{M}$ in E the hypothesis of Theorem 2 (or Theorem 2, corollary) is valid? Since, by Theorem 1, the hypothesis of Theorem 2 is valid if there exists a projection of E on $\mathfrak{M}$, and since such a projection will always exist, for every $\mathfrak{M}$, if E is finite dimensional, we see that this does not imply anything about property A, which may or may not be present (see example in §4). The projection may even be of norm 1 and yet property A be missing. This happens, for example, if E is a two-dimensional space, with $x = (x_1, x_2)$, $\| x \| = | x_1 | + | x_2 |$, for here the unit sphere has

corners.[13] On the other hand, property A is enjoyed by the reflexive spaces L^p, l^p, $(p > 1)$, and yet these contain closed linear manifolds on which there exists no projection.[14] Thus by Theorem 3 the assumptions of Theorem 2 cannot hold for all closed linear manifolds in L^p, l^p.

In order to see more precisely what *can* be said let us, for a fixed $\mathfrak{M}$, consider the operation T defined in §2. Denote by $\mathfrak{F}$ the set of elements $f \epsilon E^*$ for which $\| Tf \| = \| f \|$. $\mathfrak{F}$ is not empty, and if $\mathfrak{M}$ contains elements other than the zero element, so does $\mathfrak{F}$, by the Hahn-Banach theorem (see footnote 1). Moreover, $T(\mathfrak{F}) = \mathfrak{M}^*$, and if $f \epsilon \mathfrak{F}$, $af \epsilon \mathfrak{F}$ also, for any constant a. Now if there is a unique way of extending a linear functional from $\mathfrak{M}$ to E without increasing its norm, the totality of elements of E^* generated by this extension process is seen to be exactly $\mathfrak{F}$. Consequently, if in addition to the uniqueness of extension we suppose that the operation A of Theorem 2 exists, with the property $\| A\varphi \| = \| \varphi \|$, we conclude that $A(\mathfrak{M}^*) = \mathfrak{F}$, and hence that $\mathfrak{F}$ is a closed linear manifold in E^*. Conversely, if $\mathfrak{F}$ is a closed linear manifold in E^*, we have

THEOREM 8. *If the set $\mathfrak{F}$ is a closed linear manifold in E^*, the following things are true:*

(1) *T defines an equivalence of $\mathfrak{F}$ and $\mathfrak{M}^*$;*
(2) *each linear functional φ on $\mathfrak{M}$ has a unique extension of norm $\| \varphi \|$;*
(3) *there exists a projection, of norm* 1, *of E^* on $\mathfrak{F}$.*

Proof. (3) is a consequence of (1), by Theorem 2, corollary. (2) is likewise a consequence of (1). (1) itself is immediate if we note that $Tf = Tg$ implies $0 = \| Tf - Tg \| = \| T(f - g) \| = \| f - g \|$, whence $f = g$. Hence T maps $\mathfrak{F}$ on $\mathfrak{M}^*$ in a linear, one-to-one, isometric fashion.

The hypothesis of Theorem 8 is satisfied, for arbitrary $\mathfrak{M}$, if E is an n-dimensional Euclidean space, a Hilbert space, or any space whose norm is derived from a positive definite bilinear functional. One might suspect that this is a characteristic property of such spaces.

6. Examples. We have already quoted the fact (see footnote 14) that the reflexive spaces l^p, L^p, $(p > 1)$ contain closed linear subspaces on which there exists no projection. Since the closed subspaces of a reflexive space are also reflexive, it follows, by Theorem 3, that a subspace of L^p (or l^p) on which there exists no projection cannot satisfy the conditions imposed on the linear manifold $\mathfrak{M}$ in Theorem 2.

We next offer an example to show the need for the assumption, in Theorem 3,

[13] Any two-dimensional normed linear space always admits projections of norm 1 on an arbitrary one-dimensional subspace. See Bohnenblust, Annals of Mathematics, vol. 39(1938), p. 302, Theorem 1.

[14] L^p and l^p $(p > 1)$ have property B (see p. 79 of the reference to S. Mazur in footnote 10). Since they are reflexive they also have property A. For the assertion about linear manifolds with no projections see F. J. Murray, Transactions of the American Mathematical Society, vol. 41(1937), p. 138.

that $\mathfrak{M}$ is a reflexive subspace. Our example, in which $\mathfrak{M}$ satisfies all the conditions of the theorem except that of reflexivity, depends on the fact that for the operation A which we consider, $\| A \| = 1$, while any projection of E on $\mathfrak{M}$ must necessarily have norm $\geqq 2$. Theorem 3 asserts the existence of a projection P of norm $\| P \| = \| A \|$. It would be desirable to give an example in which there is *no* projection of E on $\mathfrak{M}$. It is an open question whether this is true or not in the example given.

For $\mathfrak{M}$ we choose the space (c_0) of sequences $x = (x_1, x_2, \cdots)$ converging to zero, with $\| x \| = \max_i | x_i |$. A linear functional on (c_0) has the form

$$\varphi(x) = \sum_{i=1}^{\infty} a_i x_i, \tag{6.1}$$

where

$$\| \varphi \| = \sum_{i=1}^{\infty} | a_i | < \infty.$$

Thus $\mathfrak{M}^*$ is equivalent to the space (l) of absolutely convergent series.

As E we take the space (m) of bounded sequences, with $\| x \| = \sup | x_i |$. Then $\mathfrak{M}$ is a closed subspace of E. The operation A of Theorem 2 obviously exists. We need only define $A\varphi = f$, where $f(x)$ is also given by (6.1), for all $x \in (m)$. Then $\| A\varphi \| = \| \varphi \|$ and $\| A \| = 1$. To see that a projection of (m) on (c_0) (if such a projection exists) must be of norm not less than 2, it suffices to consider projections of (c) on (c_0), where (c), the space of convergent sequences, is a subspace of (m), but contains (c_0) as a subspace.

Now a linear functional on (c) has the form

$$\psi(x) = \sum_{i=1}^{\infty} a_i x_i + b \lim_{n \to \infty} x_n .$$

Therefore a projection of P of (c) on (c_0) must be defined by a sequence of functionals $Px = \{\psi_i(x)\}$ of the form

$$\psi_i(x) = x_i + b_i \lim_{n \to \infty} x_n ,$$

where, since of necessity $| \psi_i(x) | \to 0$ for each x, we must have $b_i \to -1$. Moreover, $\| \psi_i \| = 1 + | b_i |$.[15] Now choose i so that $\| \psi_i \| = 1 + | b_i | > 2 - \epsilon$; next choose x, $\| x \| = 1$ so that $| \psi_i(x) | > \| \psi_i \| - \epsilon > 2 - 2\epsilon$. Thus $\| Px \| = \max_k | \psi_k(x) | > 2 - 2\epsilon$, and $\| P \| \geqq 2$, as we wished to prove.

UNIVERSITY OF CALIFORNIA AT LOS ANGELES.

[15] Banach, op. cit., p. 66–67.

CERTAIN CONGRUENCES INVOLVING THE BERNOULLI NUMBERS

By H. S. Vandiver

In a previous paper[1] I gave a general theorem concerning congruences in rings which yields congruences involving Bernoulli numbers, in particular, the relation known as Kummer's Congruence:[2]

$$h^n(h^{p-1} - 1)^j \equiv 0 \pmod{p^j} \tag{1}$$

for $n - 1 \geqq j$, $n \not\equiv 0 \pmod{p-1}$, where p is an odd prime, the left member is expanded in full by the binomial theorem and then b_t/t is substituted for h^t. The b's are defined by the recursion formula

$$(b + 1)^n = b_n \qquad (n > 1),$$

where we expand the left member by the binomial theorem and substitute b_k for b^k. The b's give the Bernoulli numbers.

A number of proofs of (1) have been given, all, as far as I know, including the restriction $n \not\equiv 0 \pmod{p-1}$; in fact, a simple inspection shows that the result does not hold when $n \equiv 0$. The question naturally arises whether there is some complementary theorem which provides for the case $n \equiv 0 \pmod{p-1}$. Nielsen[3] investigated this problem and found the relation

$$1 - \frac{1}{p} \equiv \sum_{s=1}^{s=r} (-1)^{s+s\mu} \binom{r}{s} B_{\mu s} \pmod{p},$$

$$B_n = (-1)^{n-1} b_{2n}; \qquad \mu = \tfrac{1}{2}(p-1).$$

To obtain this he sets

$$S_n(m) = 1^n + 2^n + \cdots + (m-1)^n$$

and employs the following relation (proved by him previously),

$$\sum_{k=0}^{k=r} (-1)^k \binom{r}{k} S_{2n+2k\mu}(p) \equiv \sum_{s=1}^{p-1} S^{2n}(1 - S^{2\mu})^r,$$

and

$$S_{2m}(p) \equiv (-1)^{m-1} B_m p \pmod{p^2} \qquad (p > 3,\ m > 1).$$

The result follows. Now it is not clear how this argument may be extended to the examination of (2), modulo p^a, in lieu of modulo p. But the method of my

Received January 30, 1939.

[1] Bull. Amer. Math. Soc., vol. 43(1937), pp. 417–423.

[2] Journal für die Mathematik, vol. 41(1851), pp. 368–372.

[3] *Traité des Nombres de Bernoulli*, pp. 277–278.

paper referred to in footnote 1 for the derivative of the relation given at the top of page 422 of that paper yields an extension, after certain devices are introduced. I shall confine myself to the proof of the

THEOREM. *If p is an odd prime, then*

$$b^{a(p-1)}(b^{p-1}-1)^j \equiv 0 \pmod{p^{j-1}}, \tag{2}$$

where the left member is expanded in full, b_k is substituted for b^k, and $a > 0, j > 0$, $a + j < p - 1$.

For proof we take the relation[4]

$$\frac{(n^i-1)S_i(p^k)}{p^k} = \sum_{a=1}^{p^k-1}\sum_{s=1}^{i} a^i C_{s,i}\left(\frac{y_a}{a}\right)^s p^{k(s-1)},$$

where n is prime to p and

$$y_a \equiv -\frac{a}{p} \pmod n \qquad (0 \leqq y_a < n);$$

as well as the known congruence

$$1^i + 2^i + \cdots + (p^k-1)^i \equiv p^k b_i \pmod{p^{2k}}$$

for i even and $p > 3$, the latter condition holding in the theorem since $a + j < p - 1$. These give

$$\frac{(n^{2i}-1)b_{2i}}{2i} \equiv \sum_{a=1}^{p^k-1} y_a a^{2i-1} \pmod{p^k}.$$

In the main theorem of my previous paper set $s = 2$,

$$f_{n_1} = \sum_{a=1}^{p^j-1} y_a a^{n-1},$$

$f_{n_2} = 1, \beta_1 = 1, \beta_2 = -1$, and we obtain

$$t^c(t^{p-1}-1)^j \equiv 0 \pmod{p^j}, \tag{3}$$

where

$$t_i = \frac{n^i-1}{i}\, b_i,$$

and t_i is substituted for t^i after the complete expansion of (3). For c not divisible by $(p-1)$ we may select n to be a primitive root of p and in addition such that $n^{p-1} \equiv 1 \pmod{p^j}$. Hence

$$n^{c+a(p-1)} - 1 \equiv n^c - 1 \pmod{p^j},$$

[4] Given without reference in my other paper (see footnote 1). It may be obtained from the relation (5) in the writer's paper in the Annals of Math., vol. 18(1917), p. 111, by setting $m = p^a$ and $S = 1^i + 2^i + \cdots + (p^k-1)^i$.

and after dividing by $n^c - 1$ ($\not\equiv 0 \pmod{p}$), (3) reduces to (1). However, for c a multiple of $(p-1)$ we note the known result that

$$x^{p-1} - (1 + kp) \equiv 0 \pmod{p^j}$$

always has a solution prime to p for any integer k, and set

$$n_k^{p-1} \equiv 1 + kp \pmod{p^j};$$

then

$$\frac{n_k^{r(p-1)} - 1}{p} \equiv rk + \binom{r}{2} k^2 p + \cdots + k^r p^{r-1} \pmod{p^{j-1}}.$$

Now insert this value in (3), after dividing through by p, and then divide by k. We find that

$$\frac{n_k^{r(p-1)} - 1}{pr(p-1)}$$

has been replaced by

$$\frac{1}{p-1} + \frac{1}{r(p-1)} \binom{r}{2} kp + \cdots + \frac{1}{r(p-1)} k^{r-1} p^{r-1},$$

modulo p^{j-1}. Collecting the powers of k in (3) thus obtained, we have

$$A_0 + A_1 pk + \cdots + A_{r-1} k^{r-1} p^{r-1} \equiv 0 \pmod{p^{j-1}}, \tag{4}$$

where

$$A_0 = b^c (b^{p-1} - 1)^j. \tag{5}$$

If we set $k = 1, 2, \cdots, r$ in turn, we obtain, since $r - 1 < p - 1$, a set of r congruences, and the determinant of the coefficients of the expressions $A_s p^s$ is an alternant consisting of products of the form $(i - j)$ $(i < r, j < r)$, where $i \neq j$, and each of these is, in absolute value, $< p$. Hence

$$A_0 \equiv pA_1 \equiv \cdots \equiv p^{r-1} A_{r-1} \equiv 0 \pmod{p^{j-1}}, \tag{6}$$

and with (5) this gives our theorem since $c \equiv 0 \pmod{p-1}$.

It is obvious from (6) that we have also a set of congruences moduli p^{j-2}, $p^{j-3}, \cdots, p^{j-r}$ in turn.

At first glance it may appear that (2) is less general than (1) because, at least, the modulus is p^{j-1} instead of p^j. But all the denominators of fractions occurring in (2) are divisible by p, so that if we multiply (2) through by p, we obtain by the von Staudt-Clausen theorem an expression on the left involving only fractions whose denominators are prime to p and the modulus is p^j, and this is analogous to (1).

The theorem was subject to the restriction $a > 0$. A modification of the above method will yield a theorem for $a = 0$.

It is possible to extend these ideas so that we may set up a congruence of the type given at the top of p. 422 of my previous paper with some of the n's zero,

and if $n_i = 0$, we substitute b_t for h_i^t in the result and in lieu of p^j in the modular system on the right we reduce the exponent j by one for each n_i that is zero. However, in a recent paper[5] the Bernoulli, Euler and Genocchi numbers have been generalized and some congruences of the type just mentioned may be shown to hold involving these generalized numbers; so I shall not attempt to give the most general form for our theorem until we can see how far these ideas as to the generalized numbers extend.

UNIVERSITY OF TEXAS.

[5] Vandiver, Proc. Natl. Acad., vol. 23(1937), pp. 555–559.

THE INVARIANT THEORY OF THE TERNARY TRILINEAR FORM

By Josephine H. Chanler

1. **Introduction.** Let x, y, z be three digredient ternary variables represented by points x, y, z in three respective planes E_x, E_y, E_z. We consider the trilinear form $F(x, y, z) = (\alpha x)(\beta y)(\gamma z) = \sum_{h,i,j=1}^{3} a_{hij}x_h y_i z_j$, where the a_{hij} are arbitrary complex numbers, and where we suppose that the form can be expressed in no fewer variables. The form has 27 coefficients, or 26 projective constants, of which we may eliminate $3 \cdot 8$ by proper projective transformations in E_x, E_y, E_z. For a form thus involving only 2 absolute projective constants, a complete discussion and classification of types may be anticipated.

In a joint article by R. M. Thrall and the author [12][1] the classification of such forms was given under (i) non-singular linear transformations on the sets of variables taken separately, and (ii) interchanges of the sets of variables. Two forms equivalent under (i) were called equivalent; if equivalent under (i) and (ii), they were called generally or g-equivalent. In a previous article [11] Thrall classified analogous forms in a $GF(p)$. For other references to recent studies involving the group-theoretic point of view [12] and [11] may be consulted. In [12] the approach was geometric; we studied the cubics $X(x) = 0$, $Y(y) = 0$, $Z(z) = 0$ obtained by equating to zero the determinants $|M_x(x)|$, $|M_y(y)|$, $|M_z(z)|$, where, e.g., $M_x(x)$ is the matrix $(\sum_h a_{hij}x_h)$. Particular attention was paid to the cases where the cubics had singular points or degenerated; for all such cases canonical forms for F were obtained. While the case of the elliptic cubic was discussed briefly from the analytic viewpoint, the method of attack made it impractical to set up a canonical form. In the present paper such a form is determined algebraically (see §4).

A promising method of investigation is the study of the concomitants of F. While certain of the concomitants have been previously studied, they have appeared rather in connection with bilinear than with trilinear forms. For example, C. Jordan [4] investigates the reducibility of the linear system of bilinear forms given by the trilinear form $T_{lmn} = \sum a_{\alpha\beta\gamma}\lambda_\alpha\mu_\beta x_\gamma$. For $l = m = n = 3$, he bases his discussion of the form T_{lmn} on the position of λ with respect to the cubic in λ given by the discriminant of the net of bilinear forms in μ, x. He does not mention the fact that this cubic is only one of three corresponding to our $X(x) = 0$, etc., which play equally important rôles so far as T_{lmn} is concerned. On the other hand, when J. Rosanes [8] studies the 3

Received January 30, 1939.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

ternary bilinear forms $f(x, y)$, $f^1(x, y)$, $f^2(x, y)$, he gets two cubics, $D(x^3) = 0$, $\Delta(y^3) = 0$, which are the locus of the ∞^1 common null pairs of the 3 bilinear forms. In a later article [9] Rosanes obtains a determinant invariant D whose vanishing is the condition that the 3 bilinear forms can each be expressed as a linear combination of the same 4 special (factorable) forms. He also proves that the vanishing of D insures that the 3 forms can be sent by a linear transformation upon one set of variables into 3 symmetric forms. When B. Igel [3] studies the 3 ternary bilinear forms, he considers $D(x^3) = 0, \Delta(y^3) = 0$ as together providing an analogue to the Jacobian curve of 3 quadratic forms. Maennchen [5] proves the equivalence (except for sign) of D with 2 other invariants of the net. The fact that it is an invariant of the trilinear form is used for the first time by Pasch [7] in proving the irreducibility of D. When D is so recognized, it follows readily from Rosanes' results that when $D = 0$, one can send $F = (\alpha x)(\beta y)(\gamma z)$ into the polarized form of a ternary cubic by proper projective transformations in the three planes. This theorem I have not found in the literature.

In §2 of the present article we set up the complete system of integral concomitants of degrees 1, 2, and 3 for the form F. In §3 we give their geometric interpretation where possible; the curves $X(x) = 0$, $Y(y) = 0$, $Z(z) = 0$ are related more closely to the nets of quadratic transformations by means of the concomitants. In §5 important concomitants for the canonical form derived in §4 are determined, including the invariant D above mentioned. Two absolute constants for the form are obtained. In §6 we consider cases where the transformations set up between the cubics $X(x) = 0$, $Y(y) = 0$, $Z(z) = 0$ are periodic. The invariant conditions for the early cases are found, and interesting connections are made with certain points on the cubic, such as the sextactic points and the vertices of the in-and-escribed Hart triangles.

2. **Concomitants of degrees** 1, 2, 3 **for the form** $F = (\alpha x)(\beta y)(\gamma z)$. We think of F as undergoing linear transformations on the sets of variables taken separately. Hence the concomitants, when symbolically expressed, will contain only factors of the types

$$(\xi x), \quad (\eta y), \quad (\zeta z), \quad (\alpha x), \quad (\beta y), \quad (\gamma z), \quad (\alpha_1\alpha_2\xi),$$
$$(\beta_1\beta_2\eta), \quad (\gamma_1\gamma_2\zeta), \quad (\alpha_1\alpha_2\alpha_3), \quad (\beta_1\beta_2\beta_3), \quad (\gamma_1\gamma_2\gamma_3),$$

where subscripts are used for degrees greater than 1.

For degrees 1 and 2 we have the following, besides combinations with (ξx), (ηy), (ζz):

Degree 1: $F = (\alpha x)(\beta y)(\gamma z)$ with 27 coefficients.
Degree 2: F^2, $\quad Q_\alpha = (\beta_1\beta_2\eta)(\gamma_1\gamma_2\zeta)(\alpha_1 x)(\alpha_2 x)$,

$$Q_\beta = (\alpha_1\alpha_2\xi)(\gamma_1\gamma_2\zeta)(\beta_1 y)(\beta_2 y), \quad Q_\gamma = (\alpha_1\alpha_2\xi)(\beta_1\beta_2\eta)(\gamma_1 z)(\gamma_2 z).$$

To check the completeness for degree 2 we note that F^2 has $6 \cdot 6 \cdot 6 = 216$ coefficients, while each Q has $3 \cdot 3 \cdot 6 = 54$. Thus we have a total of 378 which is the number of linearly independent quadratic combinations of the 27 coefficients of F.

For degree 3 we give the possible forms explicitly because of the syzygies to follow. Besides combinations with the absolute invariants (ξx), (ηy), (ζz) we have:

$$
\begin{aligned}
F^3 &= (\alpha_1 x)(\alpha_2 x)(\alpha_3 x)(\beta_1 y)(\beta_2 y)(\beta_3 y)(\gamma_1 z)(\gamma_2 z)(\gamma_3 z),\\
Q_\alpha \cdot F &= (\beta_1\beta_2\eta)(\gamma_1\gamma_2\zeta)(\alpha_1 x)(\alpha_2 x)(\alpha_3 x)(\beta_3 y)(\gamma_3 z),\\
Q_\beta \cdot F &= (\alpha_1\alpha_2\xi)(\gamma_1\gamma_2\zeta)(\beta_1 y)(\beta_2 y)(\alpha_3 x)(\beta_3 y)(\gamma_3 z),\\
Q_\gamma \cdot F &= (\alpha_1\alpha_2\xi)(\beta_1\beta_2\eta)(\gamma_1 z)(\gamma_2 z)(\alpha_3 x)(\beta_3 y)(\gamma_3 z),\\
B_\alpha &= (\beta_1\beta_2\beta_3)(\gamma_1\gamma_2\gamma_3)(\alpha_1 x)(\alpha_2 x)(\alpha_3 x),\\
B_\beta &= (\gamma_1\gamma_2\gamma_3)(\alpha_1\alpha_2\alpha_3)(\beta_1 y)(\beta_2 y)(\beta_3 y),\\
B_\gamma &= (\alpha_1\alpha_2\alpha_3)(\beta_1\beta_2\beta_3)(\gamma_1 z)(\gamma_2 z)(\gamma_3 z),\\
C_{\alpha\beta} &= (\alpha_1\alpha_2\alpha_3)(\beta_1\beta_2\eta)(\beta_3 y)(\gamma_1 z)(\gamma_2 z)(\gamma_3 z),\\
C_{\beta\alpha} &= (\beta_1\beta_2\beta_3)(\alpha_1\alpha_2\xi)(\alpha_3 x)(\gamma_1 z)(\gamma_2 z)(\gamma_3 z),\\
C_{\alpha\gamma} &= (\alpha_1\alpha_2\alpha_3)(\gamma_1\gamma_2\zeta)(\gamma_3 z)(\beta_1 y)(\beta_2 y)(\beta_3 y),\\
C_{\gamma\alpha} &= (\gamma_1\gamma_2\gamma_3)(\alpha_1\alpha_2\xi)(\alpha_3 x)(\beta_1 y)(\beta_2 y)(\beta_3 y),\\
C_{\beta\gamma} &= (\beta_1\beta_2\beta_3)(\gamma_1\gamma_2\zeta)(\gamma_3 z)(\alpha_1 x)(\alpha_2 x)(\alpha_3 x),\\
C_{\gamma\beta} &= (\gamma_1\gamma_2\gamma_3)(\beta_1\beta_2\eta)(\beta_3 y)(\alpha_1 x)(\alpha_2 x)(\alpha_3 x),\\
D_\alpha &= (\alpha_1\alpha_2\alpha_3)(\gamma_1\gamma_2\zeta)(\beta_1\beta_3\eta)(\beta_2 y)(\gamma_3 z),\\
D_\beta &= (\beta_1\beta_2\beta_3)(\alpha_1\alpha_2\xi)(\gamma_1\gamma_3\zeta)(\gamma_2 z)(\alpha_3 x),\\
D_\gamma &= (\gamma_1\gamma_2\gamma_3)(\beta_1\beta_2\eta)(\alpha_1\alpha_3\xi)(\alpha_2 x)(\beta_3 y),\\
E_\alpha &= (\alpha_1\alpha_2\xi)(\beta_1\beta_3\eta)(\gamma_1\gamma_3\zeta)(\alpha_3 x)(\beta_2 y)(\gamma_2 z),\\
E_\beta &= (\beta_1\beta_2\eta)(\gamma_1\gamma_3\zeta)(\alpha_1\alpha_3\xi)(\beta_3 y)(\gamma_2 z)(\alpha_2 x),\\
E_\gamma &= (\gamma_1\gamma_2\zeta)(\alpha_1\alpha_3\xi)(\beta_1\beta_3\eta)(\gamma_3 z)(\alpha_2 x)(\beta_2 y),\\
G_\alpha &= (\beta_1\beta_2\eta)(\gamma_1\gamma_3\zeta)(\beta_3 y)(\gamma_2 z)(\alpha_1 x)(\alpha_2 x)(\alpha_3 x),\\
G_\beta &= (\gamma_1\gamma_2\zeta)(\alpha_1\alpha_3\xi)(\gamma_3 z)(\alpha_2 x)(\beta_1 y)(\beta_2 y)(\beta_3 y),\\
G_\gamma &= (\alpha_1\alpha_2\xi)(\beta_1\beta_3\eta)(\alpha_3 x)(\beta_2 y)(\gamma_1 z)(\gamma_2 z)(\gamma_3 z),\\
H &= (\alpha_1\alpha_2\xi)(\beta_1\beta_3\eta)(\gamma_2\gamma_3\zeta)(\alpha_3 x)(\beta_2 y)(\gamma_1 z).
\end{aligned}
$$

These forms are related by the following syzygies:

$$H + E_\alpha - E_\gamma + D_\beta \cdot (\eta y) \equiv 0, \qquad H + E_\beta - E_\alpha + D_\gamma \cdot (\zeta z) \equiv 0,$$

$$H + E_\gamma - E_\beta + D_\alpha \cdot (\xi x) \equiv 0,$$

$$3C_{\alpha\beta} - B_\gamma \cdot (\eta y) \equiv 0, \qquad 3C_{\beta\alpha} - B_\gamma \cdot (\xi x) \equiv 0, \qquad 3C_{\alpha\gamma} - B_\beta \cdot (\zeta z) \equiv 0,$$

$$3C_{\gamma\alpha} - B_\beta \cdot (\xi x) \equiv 0, \qquad 3C_{\beta\gamma} - B_\alpha \cdot (\zeta z) \equiv 0, \qquad 3C_{\gamma\beta} - B_\alpha \cdot (\eta y) \equiv 0,$$

$$6G_\alpha - 3Q_\alpha \cdot F + B_\alpha \cdot (\eta y) \cdot (\zeta z) \equiv 0, \qquad 6G_\beta - 3Q_\beta \cdot F + B_\beta \cdot (\xi x) \cdot (\zeta z) \equiv 0,$$

$$6G_\gamma - 3Q_\gamma \cdot F + B_\gamma \cdot (\xi x) \cdot (\eta y) \equiv 0.$$

Hence we may represent all concomitants of degree 3 as linear combinations of

(1) $\quad F^3,\ Q_\alpha \cdot F,\ Q_\beta \cdot F,\ Q_\gamma \cdot F,\ B_\alpha,\ B_\beta,\ B_\gamma,\ D_\alpha,\ D_\beta,\ D_\gamma,\ E_\alpha.$

The number of coefficients for each of these is respectively 1000, 810, 810, 810, 10, 10, 10, 81, 81, 81, 729. The total is 4432 whereas the number of linearly independent cubic combinations of the coefficients of F is $C_{26+3,3} = 3654$. However, the apparent discrepancy is due to the fact that we have not chosen normal forms for our fundamental system. If we make use of a theorem of Study's ([10], p. 55) concerning the expansion of a connex $(BX)^m(UP)^n$ in a series of powers of (UX), we can prove that the number of linearly independent coefficients in our system (1) is precisely 3654. Thus (1) constitutes a complete system of concomitants of degree 3 and no further syzygies occur for that degree. The great difficulty in obtaining normal forms is due to the lack of a unique expansion for a form in more than one pair of contragredient variables ([10], p. 83).

3. **Geometric interpretations.** For given $x = x_0$,

$$(\alpha x)(\beta y)(\gamma z) = 0 \tag{1}$$

is the equation of a correlation between the planes E_y and E_z. If to points y_0, y_1 in E_y there correspond respectively the lines in E_z,

$$(\zeta_0 z) = (\alpha x_0)(\beta y_0)(\gamma z) = 0, \qquad (\zeta_1 z) = (\alpha x_0)(\beta y_1)(\gamma z) = 0,$$

then for any arbitrary line ζ on $\zeta_0\zeta_1$,

$$(\zeta_0\zeta_1\zeta) = \tfrac{1}{2}(\gamma_1\gamma_2\zeta)(\beta_1\beta_2\overline{y_0 y_1})(\alpha_1 x_0)(\alpha_2 x_0) = 0.$$

Putting $\eta = \overline{y_0 y_1}$ in E_y, we have for $x = x_0$

$$Q_\alpha = (\beta_1\beta_2\eta)(\gamma_1\gamma_2\zeta)(\alpha_1 x)(\alpha_2 x) = 0 \tag{2}$$

as the dual form of the correlation (1). If y, z are in turn taken as given constants, (1) represents in turn correlations between E_x and E_z and between E_x and E_y for which the dual forms are given by Q_β and Q_γ, respectively.

If $x = x_0$ in (1) and (2), and if y_2 is a third point in E_y corresponding to $(\zeta_2 z) = 0$ in E_z, the 3 lines $(\zeta_0 z) = 0$, $(\zeta_1 z) = 0$, $(\zeta_2 z) = 0$ are concurrent if

$$(\zeta_0\zeta_1\zeta_2) = \tfrac{1}{6}(y_0 y_1 y_2)\cdot(\beta_1\beta_2\beta_3)(\gamma_1\gamma_2\gamma_3)(\alpha_1 x_0)(\alpha_2 x_0)(\alpha_3 x_0) = 0.$$

Thus ordinarily the 3 lines ζ are concurrent only if the 3 points y are collinear. If, however, x_0 is on the cubic in E_x given by

$$X(x) = \tfrac{1}{6}B_\alpha = \begin{vmatrix} (\alpha_1 x)\beta_{11}\gamma_{11} & (\alpha_1 x)\beta_{11}\gamma_{12} & (\alpha_1 x)\beta_{11}\gamma_{13} \\ (\alpha_2 x)\beta_{22}\gamma_{21} & (\alpha_2 x)\beta_{22}\gamma_{22} & (\alpha_2 x)\beta_{22}\gamma_{23} \\ (\alpha_3 x)\beta_{33}\gamma_{31} & (\alpha_3 x)\beta_{33}\gamma_{32} & (\alpha_3 x)\beta_{33}\gamma_{33} \end{vmatrix} = 0, \tag{3}$$

the three lines are on a point z_0, whatever be y_2; that is, $(\alpha x_0)(\beta y)(\gamma z_0) \equiv 0$ in y. We say that z_0 is the singular point of the correlation in E_z, and because of the symmetry there is also a singular point y_0 in E_y. The correlations $(\alpha x)(\beta y_0)(\gamma z) = 0$ and $(\alpha x)(\beta y)(\gamma z_0) = 0$ are also singular with x_0 as a singular

point. The corresponding cubics $Y(y) = 0$, $Z(z) = 0$ are given by B_β and B_γ; and we have

THEOREM 1. *The cubics* $X(x) = 0$, $Y(y) = 0$, $Z(z) = 0$ *given by* B_α, B_β, B_γ, *respectively, are such that each gives in its plane the locus of points for which the correlation* (1) *is singular.*

If $x = x_0$ is on $X(x) = 0$, (2) becomes the product of the singular points in E_y and E_z, or

$$(2a) \qquad (\beta_1\beta_2\eta)(\gamma_1\gamma_2\zeta)(\alpha_1x_0)(\alpha_2x_0) = \rho(\eta y_0)\cdot(\zeta z_0).$$

Thus a transformation ${}_xT_y$ is set up between $X(x) = 0$ and $Y(y) = 0$, by which y_0 on $Y(y) = 0$ corresponds to x_0 on $X(x) = 0$ if y_0 is the singular point in E_y of the singular correlation determined by x_0. Also a transformation ${}_xT_z$ is set up between $X(x) = 0$ and $Z(z) = 0$, by which z_0 on $Z(z) = 0$ corresponds to x_0 on $X(x) = 0$ if z_0 is the singular point in E_z of the singular correlation determined by x_0. Analogous transformations ${}_yT_z$, ${}_yT_x$ and ${}_zT_x$, ${}_zT_y$ are given by forms derived from Q_β, Q_γ as (2a) was derived from Q_α. By the remark preceding Theorem 1, if x_0, y_0 correspond under ${}_xT_y$, they also correspond under ${}_yT_x$.

If in (2) we fix $\eta = \eta_0$,

$$(2b) \qquad Q_{xz} = (\beta_1\beta_2\eta_0)(\gamma_1\gamma_2\zeta)(\alpha_1x)(\alpha_2x) = 0$$

gives us a quadratic transformation carrying E_x into E_z in such a way that $X(x) = 0$ goes into $Z(z) = 0$ as under ${}_xT_z$. The 3 direct F-points on $X(x) = 0$ correspond to the 3 points y_0 in which η_0 meets $Y(y) = 0$. If in (2) we fix $\zeta = \zeta_0$,

$$(2c) \qquad Q_{xy} = (\beta_1\beta_2\eta)(\gamma_1\gamma_2\zeta_0)(\alpha_1x)(\alpha_2x) = 0$$

gives us a quadratic transformation carrying E_x into E_y so that $X(x) = 0$ goes into $Y(y) = 0$ as under ${}_xT_y$ and so that the 3 F-points on $X(x) = 0$ correspond to the points z_0 in which ζ_0 meets $Z(z) = 0$. Hence (2) gives us a system of conics $Q(\eta, \zeta)$ which cut out two g_2^3's coresidual to each other in the g_6^6 cut out by all conics on $X(x) = 0$. Thus, for fixed η and variable ζ we have a system of conics on a fixed triad $g_2^3(\eta)$ which cut out variable triads $g_2^3(\zeta)$. A triad $g_2^3(\eta)$ and the corresponding triad of fundamental F-points in E_z constitute in Rosanes' sense a pair of polar triangles common to one pencil of the net of bilinear forms in x, z given by (1). By bordering the determinant of $X(x)$ in (3) with η's and ζ's, we get the system $Q(\eta, \zeta)$ in the form

$$(4) \qquad Q(\eta_1\zeta) = -\tfrac{1}{2}Q_\alpha = \begin{vmatrix} (\alpha_1x)\beta_{11}\gamma_{11} & (\alpha_1x)\beta_{11}\gamma_{12} & (\alpha_1x)\beta_{11}\gamma_{13} & \eta_1 \\ (\alpha_2x)\beta_{22}\gamma_{21} & (\alpha_2x)\beta_{22}\gamma_{22} & (\alpha_2x)\beta_{22}\gamma_{23} & \eta_2 \\ (\alpha_3x)\beta_{33}\gamma_{31} & (\alpha_3x)\beta_{33}\gamma_{32} & (\alpha_3x)\beta_{33}\gamma_{33} & \eta_3 \\ \zeta_1 & \zeta_2 & \zeta_3 & 0 \end{vmatrix} = 0.$$

Q_β and Q_γ give us similar systems $Q(\xi, \zeta)$ and $Q(\eta, \xi)$ in planes E_y and E_z, respectively.

4. **Canonical form for F when $X(x) = 0$, $Y(y) = 0$, $Z(z) = 0$ are elliptic.** From [12] we sum up the necessary preliminary theorems:

If one of $X(x) = 0$, $Y(y) = 0$, $Z(z) = 0$ is a non-evanescent elliptic cubic, the others must be the same. Since in this case the cubics are birationally equivalent (under quadratic transformations), they must be projectively equivalent. We may therefore, by a suitable non-singular linear transformation on z, insure that $Z(x)$ be a constant multiple of $X(x)$, and think of E_z as superimposed on E_x. The net of quadratic transformations (given in the present paper by Q_{xz}) will then transform $X(x) = 0$ into itself. For the general elliptic cubic the only such transformations are (i) the involutions interchanging the members of every pair of points collinear with some fixed point of the cubic; (ii) the products of two such involutions.

We now turn to the algebraic treatment. As in [12] we use the theorem that the classification of the forms F is abstractly identical with the classification of the matrices $M_x(x) = (\sum_h a_{hij}x_h)$ under (i) multiplication on left and right by non-singular constant matrices and (ii) non-singular linear transformations on x. In the course of the work it will be shown that for the classes of forms obtained, no equivalences are introduced by interchange of the sets of variables. Any of our matrices $M_x(x)$ can be obtained by setting up a net of quadratic transformations carrying $X(x) = | M_x(x) | = 0$ into itself. First we consider the transformation effecting on $X(x) = 0$ the involution I_Q for which a flex point Q is the center. The net of quadratic transformations is equivalent to a collineation; that is, the form $Q(\eta, \zeta)$ of §3 factors into $Q_1(\eta)\cdot Q_2(\zeta)$, where $Q_1(\eta)$ is linear in η, $Q_2(\zeta)$ is linear in ζ, and both are linear in x. Hence if $Q(\eta, \zeta)$ is taken as a bilinear form in η, ζ, its determinant is of rank no greater than 1, whatever be the value of x. Since this determinant is to within a sign the adjoint determinant of $| M_x(x) |$, we must have $X(x) \equiv 0$. This case does not fall within the range of the present paper; indeed, its canonical form was given in [12].

We next consider transformations producing the involution I_a whose center is the generic point a upon the cubic. Since a is not a flex point, the contact points of the four tangents from a to the cubic are independent, and we may take them to be $(1, 1, 1)$, $(1, 1, -1)$, $(1, -1, 1)$, $(-1, 1, 1)$. Since the reference triangle is now also upon the cubic, the curve's equation becomes

$$a_1x_1(x_2^2 - x_3^2) + a_2x_2(x_3^2 - x_1^2) + a_3x_3(x_1^2 - x_2^2) = 0, \tag{1}$$

where (a_1, a_2, a_3) is the point a. Under the involution the four contact points are fixed; also

$$x = (a_1, 0, a_3) \leftrightarrow z = (0, 1, 0), \qquad x = (a_1, a_2, 0) \leftrightarrow z = (0, 0, 1).$$

Setting up these correspondences, we get three independent bilinear forms in x and z:

$$
(2)\qquad
\begin{aligned}
x_1z_1 \qquad\qquad\qquad\qquad - \qquad\qquad x_3z_3 &= 0,\\
x_2z_2 - \qquad\qquad x_3z_3 &= 0,\\
(a_2x_3 - a_3x_2)z_1 + (a_3x_1 - a_1x_3)z_2 + (a_1x_2 - a_2x_1)z_3 &= 0.
\end{aligned}
$$

We have therefore a net of quadratic transformations $Q_{xz}(a)$, each of which effects the involution I_a on the cubic. Eliminating the z's, we get the corresponding $M_z(x)$. Finally, we consider a transformation T_P of the cubic which is the product of two involutions. If T_P is produced by a quadratic transformation at all, it is produced by the product of a transformation of type $Q_{xz}(a)$ and an automorphism of the cubic corresponding to a flex involution I_Q. Thus the matrix determined by any T_P is equivalent to the matrix determined by some I_a.

We have now proved that every class $M_z(x)$ can be represented as

$$
(3)\qquad
\begin{pmatrix}
x_1 & 0 & -x_3\\
0 & x_2 & -x_3\\
a_2x_3 - a_3x_2 & a_3x_1 - a_1x_3 & a_1x_2 - a_2x_1
\end{pmatrix}.
$$

We want next to discover what restrictions must be imposed on the a's, and under what conditions two different sets of a's give matrices in the same class. For this purpose we study the geometric configuration presented by the net of cubics $\{a_1, a_2, a_3\} = \{a\}$ given by (1) as the point a varies. These comprise the totality of cubics on the points

$$(1, 1, 1),\ (1, 1, -1),\ (1, -1, 1),\ (-1, 1, 1),\ (0, 0, 1),\ (0, 1, 0),\ (1, 0, 0).$$

The point a is the point for which the cubic $\{a\}$ is the isologue in the transformation $x_iz_i = 1$. The four tangents from a to $\{a\}$ are

$$
(4)\qquad
\begin{aligned}
(a_2 - a_3)x_1 + (a_3 - a_1)x_2 + (a_1 - a_2)x_3 &= 0,\\
(a_2 + a_3)x_1 + (-a_3 - a_1)x_2 + (-a_1 + a_2)x_3 &= 0,\\
(a_2 + a_3)x_1 + (a_3 - a_1)x_2 + (-a_1 - a_2)x_3 &= 0,\\
(-a_2 + a_3)x_1 + (a_3 + a_1)x_2 + (-a_1 - a_2)x_3 &= 0,
\end{aligned}
$$

for which the four contact points are respectively $(1, 1, 1)$, $(1, 1, -1)$, $(1, -1, 1)$, $(-1, 1, 1)$. One value for the cross-ratio of the tangents is $(a_3^2 - a_1^2)/(a_3^2 - a_2^2)$; the other five may be obtained by permutations of a_1, a_2, a_3. The polar conic of a with respect to $\{a\}$ is

$$(5)\qquad a_1^2(x_2^2 - x_3^2) + a_2^2(x_3^2 - x_1^2) + a_3^2(x_1^2 - x_2^2) = 0.$$

As a varies over the plane, the conic (5) ranges in the pencil

$$(6)\qquad \lambda(x_2^2 - x_3^2) + \mu(x_3^2 - x_1^2) = 0.$$

The conic's parameter λ/μ in (6) is precisely the cross-ratio invariant of the corresponding cubic. The three degenerate members of the pencil,

$$x_2^2 - x_3^2 = 0, \qquad x_3^2 - x_1^2 = 0, \qquad x_1^2 - x_2^2 = 0,$$

give the locus of points a for which $\{a\}$ has at least one double point. In fact for these points the cubic degenerates. Hence for our matrix $M_z(x)$ we must avoid the values a for which

$$(a_2^2 - a_3^2)(a_3^2 - a_1^2)(a_1^2 - a_2^2) = 0. \tag{7}$$

As to equivalence, we note that the $M_z(x)$'s determined respectively by the points a and b are equivalent if and only if there exists a collineation that sends $\{a\}$ into $\{b\}$ in such a way that I_a coincides with I_b. For this to occur the set of points (1, 1, 1), (1, 1, −1), (1, −1, 1), (−1, 1, 1) can at most be permuted. There are therefore 24 collineations sending $\{a\}$ into cubics with equivalent matrices, and every point a is in a set of 24 equivalent points obtained by permutations and changes of sign of the coördinates of a. Because of the method of construction of $M_z(x)$ and the involutory character of the corresponding transformations, interchanges of the sets of variables x, y, z introduce no g-equivalences. To sum up we state

THEOREM 2. *Every ternary trilinear form $F = (\alpha x)(\beta y)(\gamma z)$ for which $X(x) = 0$ is a general elliptic cubic can be represented in the canonical form*

$$\begin{aligned}\bar{F} \equiv x_1y_1z_1 + x_2y_2z_2 - x_3y_1z_3 - x_3y_2z_3 + a_2x_3y_3z_1 \\ - a_3x_2y_3z_1 + a_3x_1y_3z_2 - a_1x_3y_3z_2 + a_1x_2y_3z_3 - a_2x_1y_3z_3 .\end{aligned} \tag{8}$$

We must avoid the values a given by (7). *The cubics $X(x) = 0$ with the same absolute invariant are the isologues of points on the* 6 *conics in* (6) *whose parameters are the* 6 *possible values of the cross-ratio invariant of the cubic. If we select one of these conics and consider on it only one point in each set*

$$(a_1, a_2, a_3), \qquad (a_1, a_2, -a_3), \qquad (a_1, -a_2, a_3), \qquad (-a_1, a_2, a_3),$$

we shall obtain one and only one representative for each class of forms corresponding to the given projective class of cubics.

5. **Certain concomitants of F calculated for the canonical form; two absolute projective constants.** For $\bar{F}$ given above (§4, (8)) we have immediately

$$X(x) = \tfrac{1}{6}B_\alpha = a_1x_1(x_2^2 - x_3^2) + a_2x_2(x_3^2 - x_1^2) + a_3x_3(x_1^2 - x_2^2),$$

$$Y(y) = \tfrac{1}{6}B_\beta = (a_1^2 - a_3^2)y_1y_3^2 + (a_2^2 - a_3^2)y_2y_3^2 - y_1y_2(y_1 + y_2),$$

$$Z(z) = \tfrac{1}{6}B_\gamma = a_1z_1(z_3^2 - z_2^2) + a_2z_2(z_1^2 - z_3^2) + a_3z_3(z_2^2 - z_1^2),$$

$$Q(\eta, \zeta) = -\tfrac{1}{2}Q_\alpha = \begin{vmatrix} x_1 & 0 & -x_3 & \eta_1 \\ 0 & x_2 & -x_3 & \eta_2 \\ a_2x_3 - a_3x_2 & a_3x_1 - a_1x_3 & a_1x_2 - a_2x_1 & \eta_3 \\ \zeta_1 & \zeta_2 & \zeta_3 & 0 \end{vmatrix},$$

$$Q(\xi, \zeta) = -\tfrac{1}{2}Q_\beta = \begin{vmatrix} y_1 & a_3y_3 & -a_2y_3 & \xi_1 \\ -a_3y_3 & y_2 & a_1y_3 & \xi_2 \\ a_2y_3 & -a_1y_3 & -y_1-y_2 & \xi_3 \\ \zeta_1 & \zeta_2 & \zeta_3 & 0 \end{vmatrix},$$

$$Q(\eta, \xi) = -\tfrac{1}{2}Q_\gamma = \begin{vmatrix} z_1 & 0 & -z_3 & \eta_1 \\ 0 & z_2 & -z_3 & \eta_2 \\ a_3z_2 - a_2z_3 & a_1z_3 - a_3z_1 & a_2z_1 - a_1z_2 & \eta_3 \\ \xi_1 & \xi_2 & \xi_3 & 0 \end{vmatrix}.$$

It is of interest to calculate the quartinvariant S and the sextinvariant T for the cubic $X(x) = 0$. We use Cayley's table ([1], p. 325), but to get our S and T multiply those given by Cayley by 81 and $-\frac{729}{4}$, respectively.

$$S = a_1^4 + a_2^4 + a_3^4 - a_1^2a_2^2 - a_2^2a_3^2 - a_3^2a_1^2,$$

$$T = (2a_3^2 - a_1^2 - a_2^2)(2a_2^2 - a_3^2 - a_1^2)(2a_1^2 - a_2^2 - a_3^2).$$

We may arrive immediately at these expressions by noting that if $S = 0$, the cubic is equianharmonic with cross-ratio $(a_3^2 - a_1^2)/(a_3^2 - a_2^2) = \sigma$, where $\sigma^2 - \sigma + 1 = 0$, while if $T = 0$, the cross-ratio $(a_3^2 - a_1^2)/(a_3^2 - a_2^2) = -1, 2,$ or $\frac{1}{2}$. The discriminant is

$$R = 4S^3 - T^2 = 27[(a_1^2 - a_2^2)(a_2^2 - a_3^2)(a_3^2 - a_1^2)]^2,$$

checking the range of excluded values given by (7), §4. The cases where $R = 0$ are considered explicitly in [12]. Considered as invariants of the ternary trilinear form, these invariants have the following degrees and weights, the weights being calculated for each set of variables:

S: degree 12, weight 4;

T: degree 18, weight 6;

R: degree 36, weight 12.

Written symbolically for $F = (\alpha x)(\beta y)(\gamma z)$,

$$S = K_1(\alpha_1\alpha_2\alpha_3)(\alpha_4\alpha_5\alpha_6)(\alpha_7\alpha_8\alpha_9)(\alpha_{10}\alpha_{11}\alpha_{12})(\beta_1\beta_4\beta_7)(\beta_2\beta_5\beta_{10})(\beta_3\beta_8\beta_{11})(\beta_6\beta_9\beta_{12}) \cdot (\gamma_1\gamma_4\gamma_7)(\gamma_2\gamma_5\gamma_{10})(\gamma_3\gamma_8\gamma_{11})(\gamma_6\gamma_9\gamma_{12}),$$

$$T = K_2(\alpha_1\alpha_2\alpha_3)(\alpha_4\alpha_5\alpha_6)(\alpha_7\alpha_8\alpha_9)(\alpha_{10}\alpha_{11}\alpha_{12})(\alpha_{13}\alpha_{14}\alpha_{15})(\alpha_{16}\alpha_{17}\alpha_{18}) \cdot (\beta_1\beta_4\beta_7)(\beta_2\beta_5\beta_{10})(\beta_3\beta_8\beta_{11})(\beta_6\beta_{13}\beta_{16})(\beta_9\beta_{14}\beta_{17})(\beta_{12}\beta_{15}\beta_{18}) \cdot (\gamma_1\gamma_4\gamma_7)(\gamma_2\gamma_5\gamma_{10})(\gamma_3\gamma_8\gamma_{11})(\gamma_6\gamma_{13}\gamma_{16})(\gamma_9\gamma_{14}\gamma_{17})(\gamma_{12}\gamma_{15}\gamma_{18}),$$

where K_1, K_2 are constants. For the canonical form the invariants S, T, R are precisely the same for all the cubics. For the form $F = (\alpha x)(\beta y)(\gamma z)$ the condition $R = 0$ gives the cases of singular cubics studied in [12].

The invariant of lowest degree for $F = (\alpha x)(\beta y)(\gamma z)$ is $(\alpha_1\alpha_2\alpha_3)(\alpha_4\alpha_5\alpha_6)(\beta_1\beta_3\beta_5) \cdot (\beta_4\beta_6\beta_2)(\gamma_1\gamma_2\gamma_6)(\gamma_4\gamma_5\gamma_3)$, of degree 6 and weight 2 for each set of variables. It may be obtained as the apolarity invariant of any one of D_α, D_β, D_γ with itself. For the canonical form $\bar{F}$, it is to within a constant factor equal to

$$I_2 = a_1^2 + a_2^2 + a_3^2.$$

The invariant D mentioned in §1 is given by Pasch for the form $\sum a_{hij}x_hy_iz_j$ as a determinant of order 9. Of degree 9, it is of weight 3 for each set of variables. For the canonical form $\bar{F}$, it is to within a constant factor equal to

$$I_3 = a_1a_2a_3 ,$$

which will figure in §6.

If we add the invariant $I_4 = a_1^2a_2^2 + a_2^2a_3^2 + a_3^2a_1^2 = \frac{1}{3}(I_2^2 - S)$, of degree 12 and weight 4 for each set of variables, we may set up two absolute projective constants for the ternary trilinear form: I_4/I_2^2 and I_3^2/I_2^3. When definite values are assigned to these absolute invariants, and the resulting expressions cleared of fractions and treated as equations of the fourth and sixth degrees respectively in the homogeneous variables a_1, a_2, a_3, we get 24 independent solutions. Since the equations are symmetrical in the a's and contain their even powers only, these 24 comprise a conjugate set of a's which produce equivalent forms. Hence we have

Theorem 3. *Two ternary trilinear forms for which $X(x) = 0$ is a general elliptic cubic are equivalent if and only if they have the same absolute invariants I_4/I_2^2, I_3^2/I_2^3.*

Since the determination of a class of forms depends upon the choice of the absolute invariant of the cubic $X(x) = 0$, as well as upon the selection of the point a (and therefore the coresidual g_2^3's) upon the cubic, it would be advantageous to choose for one of the 2 constants the absolute invariant $J = S^3/T^2$. If, however, J and I_3^2/I_2^3 are taken for the constants, we get for definite values of these constants, 72 sets of a's which break up into 3 equal groups; the forms produced by a's in the same group are equivalent but not those produced by a's in different groups. Likewise the constants J and I_4/I_2^2 determine the forms only to within 2 non-equivalent classes. Nevertheless we shall often use J in the next section.

By the simultaneous transformations

$$\begin{aligned} x_1' &= x_2, & y_1' &= y_2, & z_1' &= z_2, \\ x_2' &= x_1, & y_2' &= y_1, & z_2' &= z_1, \qquad (1) \\ x_3' &= x_3, & y_3' &= -y_3, & z_3' &= z_3, \end{aligned}$$

$\bar{F}$ may be sent into a form differing from $\bar{F}$ only in the permutation of a_1, a_2. Since the weight of any (integral) invariant must be the same positive integer for the three sets of variables, any such invariant is multiplied by unity when the transformations (1) are carried out on $\bar{F}$. Hence all (integral) invariants of $\bar{F}$ are symmetric in a_1, a_2 and may be similarly proved symmetric in a_1, a_2, a_3. They are therefore integral rational functions of the elementary symmetric functions

$$s_1 = a_1 + a_2 + a_3 , \qquad s_2 = a_1a_2 + a_2a_3 + a_3a_1 , \qquad s_3 = I_3 .$$

Of these, s_1 is an algebraic invariant (cf. [10], pp. 10, 11) of degree 3 and weight 1 for each set of variables; s_2 is an algebraic invariant of degree 6 and weight 2. For F they are determined respectively as roots of the following equations:

$$s_1^8 - 4I_2s_1^6 + (6I_2^2 - 8I_4)s_1^4 + (16I_2I_4 - 4I_2^3 - 64I_3^2)s_1^2 + I_2^4 - 8I_2^2I_4 + 16I_4^2 = 0,$$

$$s_2^4 - 2I_4s_2^2 - 8I_3^2s_2 + I_4^2 - 4I_3^2I_2 = 0. \tag{2}$$

Also

$$s_2 = \tfrac{1}{2}(s_2^1 - I_2).$$

The 8 conjugate values of s_1 and the 4 conjugate values of s_2 given by equations (2) are obtained for the canonical form $\bar{F}$ by changes of sign of a_1, a_2, a_3.

If we take s_2/s_1^2, s_3/s_1^3 as the absolute constants of the form F, we get for definite values of these constants, 6 sets of a's which all give rise to equivalent forms. However, any definite pair of values for s_2/s_1^2, s_3/s_1^3 is associated with 7 other pairs which give rise to the same class of forms.

We have almost immediately the theorem (verified by Maennchen's calculations and transformations) that D or I_3 is the sole integral invariant of degree 9. For any such invariant, being of weight 3, must be a linear combination of s_1^3, $s_1^2s_2$ and $s_3 = I_3$. However, the first two forms are not integral rational invariants since any integral invariant containing s_1 as a factor must contain its conjugates and be of weight 8 at least.

6. **Periodic transformations among the cubics** $X(x) = 0$, $Y(y) = 0$, $Z(z) = 0$.

(a) *General discussion.* A consideration of Q_{xz}, Q_{xy} (§3, (2b), (2c)) leads to

THEOREM 4. *We assign the elliptic parameter u on $X(x) = 0$, v on $Y(y) = 0$, and w on $Z(z) = 0$, so that u is the canonical parameter on $X(x) = 0$, while x_0, y_0, z_0 have equal parameters $u = v = w$ if y_0, z_0 correspond to x_0 under ${}_xT_y$, ${}_xT_z$, respectively. Then on $Y(y) = 0$, $Z(z) = 0$, the collinear conditions are respectively*

$$v_1 + v_2 + v_3 \equiv k, \qquad w_1 + w_2 + w_3 \equiv -k.$$

If we consider the systems of conics $Q(\xi, \zeta)$ and $Q(\eta, \xi)$ with their corresponding sets of transformations, we have for the g_2^3's corresponding to the line sections in the various planes

$$\begin{array}{llll} & X(x) = 0, & & Y(y) = 0, \\ \text{lines } \xi & : u_1 + u_2 + u_3 \equiv 0, & g_2'^3(\xi) & : v_1 + v_2 + v_3 \equiv 0, \\ g_2^3(\eta) & : u_1 + u_2 + u_3 \equiv k, & \text{lines } \eta & : v_1 + v_2 + v_3 \equiv k, \\ g_2^3(\zeta) & : u_1 + u_2 + u_3 \equiv -k, & g_2'^3(\zeta) & : v_1 + v_2 + v_3 \equiv 2k, \end{array}$$

$$\begin{array}{ll} & Z(z) = 0, \\ g_2''^3(\xi) & : w_1 + w_2 + w_3 \equiv 0, \\ g_2''^3(\eta) & : w_1 + w_3 + w_3 \equiv -2k, \\ \text{lines } \zeta & : w_1 + w_2 + w_3 \equiv -k, \end{array}$$

where, e.g., in the first line the points of the triads $g_2'^3(\xi)$ on $Y(y) = 0$, and $g_2''^3(\xi)$ on $Z(z) = 0$ correspond to the points of the line ξ on $X(x) = 0$ under ${}_xT_y$, ${}_xT_z$, respectively. Hence if y_0, z_0 correspond under ${}_yT_z$, the relation between their parameters is

$$w \equiv v - k.$$

There follows immediately

THEOREM 5. *Let a sequence*

$$x_0y_0z_1x_1y_1 \cdots z_ix_iy_iz_{i+1} \cdots \tag{1}$$

be given, where x_i corresponds to z_i under ${}_zT_x$, y_i corresponds to x_i under ${}_xT_y$, and z_{i+1} corresponds to y_i under ${}_yT_z$. Then the necessary and sufficient condition that the sequence be periodic is that k be a submultiple of an elliptic period. The same conclusion holds for the analogous sequence

$$x_0z_0y_1x_1z_1 \cdots y_ix_iz_iy_{i+1} \cdots .$$

For (1) the corresponding sequence of transformations on the elliptic parameters is

$$v_0 = u_0, \; w_1 = u_0 - k, \; u_1 = u_0 - k, \; v_1 = u_0 - k, \; w_2 = u_0 - 2k, \; \cdots . \tag{1a}$$

We shall assume that

$$nk \equiv 0 \qquad \text{mod } (\omega_1, \omega_2), \tag{2}$$

where ω_1, ω_2 are the fundamental elliptic periods, and shall consider the transformation on $X(x) = 0$ produced by ${}_xT_y \cdot {}_yT_z \cdot {}_zT_x$. Obviously similar ones take place on the other cubics. Then in E_x we have for triads $g_2^3(\eta)$ and $g_2^3(\zeta)$, respectively,

$$u_1 + u_2 + u_3 \equiv \frac{\lambda\omega_1 + \mu\omega_2}{n}, \qquad u_4 + u_5 + u_6 \equiv -\frac{\lambda\omega_1 + \mu\omega_2}{n}, \tag{3}$$

and the series $g_2^3(\eta)$ (or the series $g_2^3(\zeta)$) is cut out by an n-ic with n-fold contact at the 3 points given by u_1, u_2, u_3 (or u_4, u_5, u_6). If $n \sum_{i=1}^{3} u_i \equiv 0$, the conics on u_1, u_2, u_3 cut out a series $g_2^3(\zeta)$ such that for each group u_4, u_5, u_6 in it, $\sum_{i=1}^{3} u_i + \sum_{i=4}^{6} u_i \equiv 0$; whence $n \sum_{i=4}^{6} u_i \equiv 0$. We thus have

THEOREM 6. *The existence on the cubic curve of one group u_1, u_2, u_3 such that $n \sum_{i=1}^{3} u_i \equiv 0$, insures the existence of two coresidual series $g_2^3(\eta)$ and $g_2^3(\zeta)$ which produce periodic transformations.*

An n-ic with n-fold contact at 3 points is subject to $3n$ conditions. However, only $3n - 1$ points can be chosen on the cubic to determine the n-ic. Hence

there is always one condition on the cubic and the set of points u_1, u_2, u_3 that the set be in a $g_2^3(\eta)$ or $g_2^3(\zeta)$ of the desired type. From analysis it is obvious that this condition can be put upon the set of points. The two coresidual $g_2^3(\eta)$ and $g_2^3(\zeta)$ are identical only when $n = 1$ or 2. In all other cases the 9 collineations of the cubic into itself given by the flex involutions: $3(u' + u) \equiv 0$ interchange the $g_2^3(\eta)$ with the $g_2^3(\zeta)$, while the 9 collineations $3(u' - u) \equiv 0$ send each series into itself.

(b) *Case* $n = 1$. Here the transformation carried out on $X(x) = 0$ by ${}_xT_y \cdot {}_yT_z \cdot {}_zT_x$ is the identity. Since the series $g_2^3(\eta)$, $g_2^3(\zeta)$ are cut out on $X(x) = 0$ by all the straight lines in E_x, the system of conics $Q(\eta, \zeta)$ factors as in §4, as do $Q(\xi, \zeta)$ and $Q(\eta, \xi)$. Hence $X(x) \equiv Y(y) \equiv Z(z) \equiv 0$. The identical vanishing of any 2 (and therefore of all 3) of B_α, B_β, B_γ constitutes the necessary and sufficient conditions for this case. We cannot set up the conditions on the a's as for the other cases, since this situation cannot occur for our canonical form.

(c) *Case* $n = 2$. Here the $g_2^3(\eta) = g_2^3(\zeta)$ and k is one of $\frac{1}{2}\omega_1$, $\frac{1}{2}\omega_2$, $\frac{1}{2}(\omega_1 + \omega_2)$; whence we have

THEOREM 7. *Given the invariant J, there exist 3 non-equivalent classes of ternary trilinear forms for which the transformation ${}_xT_y \cdot {}_yT_z \cdot {}_zT_x$ on $X(x) = 0$ is of period 2. The g_2^3's corresponding to these classes are cut out on $X(x) = 0$ by the contact conics constituting the 3 systems of poloconics of the lines of the plane with respect to the 3 cubics for which $X(x) = 0$ is the Hessian. The conics of the system $Q(\eta, \zeta)$ are the mixed poloconics for the line pairs of the plane.* (For definitions and references for these systems of conics cf. [2], pp. 471–472; [6], p. 379.)

For our canonical form let the elliptic parameter of the point a be α. Then the parameters of (1, 1, 1), (1, 1, −1), (1, −1, 1), (−1, 1, 1) are (not necessarily in this order) congruent to $-\frac{1}{2}\alpha$, $\frac{1}{2}(-\alpha + \omega_1)$, $\frac{1}{2}(-\alpha + \omega_2)$, $\frac{1}{2}(-\alpha + \omega_1 + \omega_2)$. The diagonal triangle of this four-point is the reference triangle; its vertices have parameters congruent to $\alpha + \frac{1}{2}\omega_1$, $\alpha + \frac{1}{2}\omega_2$, $\alpha + \frac{1}{2}(\omega_1 + \omega_2)$. Since these vertices constitute a triad $g_2^3(\eta)$, we must have $6\alpha \equiv 0$. There are 36 such points, of which the 9 flexes must be discarded, leaving for possible a's the 27 sextactic points which are the intersections with the cubic of the 9 harmonic polars. From the parameters it is evident that a must lie on one side of the reference triangle. Hence

THEOREM 8. *In general if $R \neq 0$, the necessary and sufficient condition that the transformation ${}_xT_y \cdot {}_yT_z \cdot {}_zT_x$ on $X(x) = 0$ be of period 2 is that $I_3 = 0$.*

When a is chosen, two vertices of the reference triangle prove to be the two other sextactic points upon the harmonic polar through a, while the third is the corresponding flex point. Thus the particular g_2^3 depends upon the choice of 2 out of 3 sextactic points to serve with their flex as a group of the g_2^3.

(d) *Case* $n = 3$. If we take α to be the parameter of the point a as in case (c), we find that here $9\alpha \equiv 0$, or a is a coincidence point (as defined in [2], p. 501). Excluding the flexes, there are 72 such points; taking account of the 18 collinea-

tions of the general cubic into itself, we should expect $\frac{72}{18} = 4$ non-equivalent a's; i.e., 4 non-equivalent pairs of coresidual g_2^3's on the cubic. Furthermore the 72 coincidence points are known to be the vertices of the 24 Hart triangles in-and-escribed to the cubic ([2], p. 501). If the vertices of such a triangle have parameters u_1, u_2, u_3, we have $3(u_1 + u_2 + u_3) \equiv 0$; hence these vertices constitute a group of $g_2^3(\eta)$ or of $g_2^3(\zeta)$ of the type described. It is known that the Hart triangles can be divided into 4 sets of 6 each such that the triangles in a set are permuted among themselves by the 18 collineations. The vertices of the 6 triangles in a set prove to be 3 groups of one $g_2^3(\eta)$ and 3 groups of its coresidual $g_2^3(\zeta)$; the 8 ternary cyclic collineations permute the triangles of each series among themselves; the 9 flex involutions interchange the triangles of $g_2^3(\eta)$ with those of $g_2^3(\zeta)$. If a has parameter α, the sum of the parameters of the vertices of the Hart triangle of which a is a vertex is congruent to 3α; this is precisely the congruence sum obtained for the group $g_2^3(\eta)$ of the reference triangle.

THEOREM 9. *Given the invariant J, there exist in general* 4 *non-equivalent classes of ternary trilinear forms for which the transformation ${}_xT_y \cdot {}_yT_z \cdot {}_zT_x$ on $X(x) = 0$ is of period* 3. *The* 4 *pairs of coresidual g_2^3's correspond to the* 4 *sets of Hart triangles; each pair of g_2^3's contains as groups the triads of vertices of one such set of* 6.

Three non-collinear flex points also constitute a group of such a $g_2^3(\eta)$ or $g_2^3(\zeta)$, as do one flex point and a second flex point taken twice. Of the flex-group for which we have one repeated, we have $2C_{9,2} = 72$; of the flex groups of 3 different points we have $C_{9,3} - 12 = 72$, since they must be non-collinear. The argument goes forward as before, giving $\frac{72}{18} = 4$ non-equivalent pairs of coresidual $g_2^3(\eta)$, $g_2^3(\zeta)$.

To determine the invariant for $n = 3$, we use the fact that the point a is its own third tangential. The first tangential is

$$x_1' = a_2a_3, \qquad x_2' = a_3a_1, \qquad x_3' = a_1a_2; \tag{4}$$

the second is

$$\begin{gathered} x_1'' = a_2a_3(a_2^2a_3^2 - a_3^2a_1^2 - a_1^2a_2^2), \qquad x_2'' = a_3a_1(a_3^2a_1^2 - a_1^2a_2^2 - a_2^2a_3^2), \\ x_3'' = a_1a_2(a_1^2a_2^2 - a_2^2a_3^2 - a_3^2a_1^2). \end{gathered} \tag{5}$$

Writing out the equation for the tangent at x'' and imposing the condition that the coördinates of a satisfy this equation, we get

THEOREM 10. *In general, if $R \neq 0$, $I_3 \neq 0$, the necessary and sufficient condition that the transformation ${}_xT_y \cdot {}_yT_z \cdot {}_zT_x$ on $X(x) = 0$ be of period* 3 *is that*

$$a_1^8(a_2^2 - a_3^2)^3 + a_2^8(a_3^2 - a_1^2)^3 + a_3^8(a_1^2 - a_2^2)^3 = 0. \tag{6}$$

As a check consider the number of independent solutions of (6) solved simultaneously with the conic in the a's for given cross-ratio λ/μ:

$$\lambda(a_2^2 - a_3^2) + \mu(a_3^2 - a_1^2) = 0. \tag{7}$$

We have $2 \cdot 14 = 28$ solutions, of which $4 \cdot 3 = 12$ must be excluded, since $(1, 1, 1)$, $(1, 1, -1)$, $(1, -1, 1)$, $(-1, 1, 1)$ are triple points of (6) and simple points of (7). Remembering that the 4 sets of values $(a_1, \pm a_2, \pm a_3)$ produce forms in the same class, we get $\frac{16}{4} = 4$ non-equivalent classes.

Bibliography

1. A. Cayley, *The Collected Mathematical Papers*, vol. 2, 1889.
2. Encyklopädie der Mathematischen Wissenschaften, vol. 3, C5.
3. B. Igel, *Ueber drei Paare von Hauptdreiecken in der Theorie der bilinearen Formen*, Monatshefte für Mathematik und Physik, vol. 5(1894), pp. 289–302.
4. C. Jordan, *Réduction d'un réseau de formes quadratiques ou bilinéaires* (deuxième partie), Jour. de Math., (6), vol. 3(1907), pp. 5–51.
5. P. Maennchen, *Die Transformation der trilinearen ternären Form*, 1898.
6. E. Pascal, *Repertorium der höheren Mathematik*, vol. 2, part 1, 1910.
7. M. Pasch, *Ueber eine Invariante der trilinearen ternären Form*, Mathematische Annalen, vol. 52(1899), pp. 127–129.
8. J. Rosanes, *Ueber linear-abhängige Punktsysteme*, Journal für die reine und angewandte Mathematik, vol. 88(1880), pp. 241–273.
9. J. Rosanes, *Ueber abhängige Punktsysteme und deren Bedeutung für die reciproke Verandtschaft zweier Ebenen*, Journal für die reine und angewandte Mathematik, vol. 95(1883), pp. 247–255.
10. E. Study, *Methoden zur Theorie der ternaeren Formen*, 1889.
11. R. M. Thrall, *Metabelian groups and trilinear forms*, American Journal of Mathematics, vol. 60(1938), pp. 383–415.
12. R. M. Thrall and J. H. Chanler, *Ternary trilinear forms in the field of complex numbers*, this Journal, vol. 4(1938), pp. 678–690.

University of Illinois.

THE PROBLEM OF TYPE FOR A CERTAIN CLASS OF RIEMANN SURFACES

By F. E. Ulrich

1. **Introduction.** In the present paper there are obtained sufficient conditions that a certain class of open simply-connected Riemann surfaces shall be of parabolic type,[1] that is, can be mapped one-to-one and in general conformally on the plane with one point removed. The surfaces to be considered are, briefly, surfaces which have a single transcendental singularity which is a limit point of algebraic branch points of first order, and a logarithmic branch point.

The results are derived from a criterion due to Ahlfors.[2] It can be stated as follows. Let the open simply-connected surface W be spread out over the w-plane and a metric on W be defined by a differential form

$$d\sigma = \lambda(u, v) \mid dw \mid, \qquad w = u + iv,$$

where λ is a real, single-valued function, continuous on W with the exception of certain isolated points. Moreover, let the metric be so chosen that no singularity of the surface is accessible along a path of finite length. Let W_ρ be the region of the surface consisting of those points whose distance from a certain initial point P_0, in the metric considered, does not exceed a positive number ρ. Let $L(\rho)$ be the length of the boundary of W_ρ in the metric considered. Then, *a necessary and sufficient condition that W be of parabolic type is that there exist a metric of the type defined above such that the integral*

$$\int^{\infty} \frac{d\rho}{L(\rho)}$$

diverges.

2. **Class of surfaces to be considered.** The surfaces W to be considered are of the following sort.

Let $\{A_\nu\}$ $(\nu = 1, 2, 3, \cdots)$ be a countably infinite set of points of the real axis of the w-plane, which has as sole limit point the point at infinity. Moreover, suppose $A_\nu > 0$ for ν odd and $A_\nu < 0$ for ν even. Over each point of $\{A_\nu\}$ shall lie one and only one algebraic branch point of first order. There shall be no other algebraic branch points. There shall be a logarithmic branch point over $w = 0$ along with an infinite number of smooth sheets. There will

Received February 3, 1939.

[1] The writer wishes to express his thanks to Professor Ahlfors for suggesting the problem and for the valuable counsel given by him.

[2] Comptes Rendus, vol. 201(1935), pp. 30-32.

be a transcendental singularity over $w = \infty$, which is a limit point of algebraic branch points.[3]

Since all branch points and singularities of W lie over points of the real axis, the sheets can be divided into half-sheets by a cut along this line. All branch points and singularities will then be on the boundaries of these half-sheets.

The half-sheets thus formed shall in general be of two sorts. The upper and lower half-sheets of the first kind shall be denoted respectively by S_ν^+ and S_ν^- ($\nu = 1, 2, 3, \cdots$) and be spoken of as the algebraic half-sheets. For each ν, both S_ν^+ and S_ν^- shall abut on the algebraic branch points over A_ν and $A_{\nu+1}$, as well as the transcendental singularity over $w = \infty$, but together form a smooth sheet over each of the remaining points of $\{A_\nu\}$ as well as over $w = 0$.

It is convenient to separate the half-sheets of the second kind in two infinite sets. The upper and lower half-sheets of the first set shall be denoted respectively by Q_ν^+ and Q_ν^- ($\nu = 1, 2, 3, \cdots$), those of the second set respectively by $Q_{-\nu}^+$ and $Q_{-\nu}^-$ ($\nu = 1, 2, 3, \cdots$). Each half-sheet of both sets shall abut on the singularities over $w = 0$ and $w = \infty$. But an upper and a lower half-sheet with the same index together shall form a smooth sheet over each point of the set $\{A_\nu\}$.

There shall be one remaining sheet, the upper and lower half-sheets of which shall be denoted respectively by Q_0^+ and Q_0^-. Both Q_0^+ and Q_0^- shall abut on the singularities over $w = 0$ and $w = \infty$ as well as the algebraic branch point over A_1, but together shall form a smooth sheet over each of the remaining points of $\{A_\nu\}$.

The half-sheets $Q_\nu^\pm$ ($\nu = 0, \pm 1, \pm 2, \cdots$) shall be spoken of as the logarithmic half-sheets.

The half-sheets are joined in the following way:

The two sets $Q_\nu^\pm$ ($\nu = \pm 1, \pm 2, \pm 3, \cdots$) form two logarithmic ends, one formed by the set $Q_\nu^\pm$ ($\nu = 1, 2, 3, \cdots$) joined to Q_0^+ along the segment $(0, -\infty)$, the other by the set $Q_\nu^\pm$ ($\nu = -1, -2, -3, \cdots$) joined to Q_0^- along $(0, -\infty)$. More completely stated, to $Q_{\nu-1}^+$ is joined Q_ν^- along the segment $(0, -\infty)$ and to Q_ν^- is joined Q_ν^+ along the segment $(0, +\infty)$ ($\nu = 1, 2, 3, \cdots$). This joining forms one of the logarithmic ends. Then to $Q_{\nu+1}^-$ is joined Q_ν^+ along $(0, -\infty)$ and to Q_ν^- is joined Q_ν^+ along $(0, +\infty)$ ($\nu = -1, -2, -3, \cdots$). This joining forms the second logarithmic end.

To complete the surface W, Q_0^+ and Q_0^- are joined along the segment $(0, A_1)$ and the algebraic half-sheets are adjoined as follows: To Q_0^- is joined S_1^+ along the segment $(A_1, +\infty)$, to Q_0^+ is joined S_1^- along $(A_1, +\infty)$ and to S_1^+ is joined S_1^- along (A_1, A_2). Now let (A_ν, ∞) ($\nu = 1, 2, 3, \cdots$) be that infinite segment which is on the real axis of the w-plane, which has finite end point A_ν, and which does not contain the point $A_{\nu+1}$. Then for $\nu \geqq 2$, to S_ν^- is joined S_ν^+ along

[3] It should be pointed out at this time that the requirement that the points A_ν have the point at infinity as a limit point is entirely unnecessary for the work to follow. But it will be seen (§11) that the sufficient conditions obtained are not met unless $|A_\nu|$ becomes infinite with increasing ν. If the $|A_\nu|$ are bounded, the character of the singularity of W over $w = \infty$ is changed. Instead of a transcendental singularity of the sort described above, there will be two logarithmic branch points.

$(A_\nu, A_{\nu+1})$, $S^+_{\nu-1}$ along (A_ν, ∞) and $S^+_{\nu+1}$ along $(A_{\nu+1}, \infty)$. To S^+_ν is joined S^-_ν along $(A_\nu, A_{\nu+1})$, $S^-_{\nu-1}$ along (A_ν, ∞) and $S^-_{\nu+1}$ along $(A_{\nu+1}, \infty)$.

The algebraic half-sheets S^+_ν and S^-_ν together shall form the algebraic sheet S_ν ($\nu = 1, 2, 3, \cdots$), and the logarithmic half-sheets Q^+_ν and Q^-_ν together shall form the logarithmic sheet Q_ν ($\nu = 0, \pm 1, \pm 2, \cdots$).

3. **The metric on** W. A metric on the surface W will be defined by the differential form

$$d\sigma = \lambda(u, v) \mid dw \mid, \qquad w = u + iv,$$

σ being arc length in the non-Euclidean metric. The function $\lambda(u, v) = \lambda(w)$ shall be a real, single-valued function, continuous on W except for isolated points. $\lambda(w)$ is defined as follows: Let $\tilde{w}$ be a place on W over the point w of the w-plane. Let $\Delta(\tilde{w})$ be the Euclidean distance *on the surface* from $\tilde{w}$ to the essential singularity over $w = 0$. Here distance is used in the sense of shortest distance, or more precisely, the greatest lower bound of all Euclidean distances on W from $\tilde{w}$ to the singular point over $w = 0$. We then define

$$\lambda(w) = \frac{1}{\Delta(\tilde{w})},$$

and have

$$d\sigma = \frac{\mid dw \mid}{\Delta(\tilde{w})}.$$

In the logarithmic sheets Q_ν, since they abut on the singularity over $w = 0$,

$$\Delta(\tilde{w}) = \mid w \mid.$$

Since the algebraic sheets S_ν do not abut on the singularity over $w = 0$, in determining $\Delta(\tilde{w})$ for $\tilde{w}$ in one of these sheets, it is necessary to follow a path on the surface from $\tilde{w}$ into one of the sheets which do abut on this singular point. From the way in which the sheets of W are connected, it is clear that a shortest path is determined by going from $\tilde{w}$ in the sheet S_ν in question, through the sheets $S_{\nu'}$ with index $\nu' < \nu$, into the sheet Q_0 and hence to the singularity. We then have for $\tilde{w}$ in S_ν,

$$\Delta(\tilde{w}) = \mid w - A_\nu \mid + \mid A_\nu - A_{\nu-1} \mid + \cdots + \mid A_2 - A_1 \mid + A_1 .$$

If we let

$$\mid A_\nu - A_{\nu-1} \mid = p_{\nu-1}, \qquad \nu \geqq 2, \qquad A_1 = p_0, \tag{1}$$

and

$$c_\nu = p_0 + p_1 + \cdots + p_{\nu-1}, \qquad \nu \geqq 1, \tag{2}$$

the analytic expression for the differential form defining the metric on W has the following form in the various sheets.

In all the logarithmic sheets Q_ν,

$$d\sigma = \frac{|\,dw\,|}{|\,w\,|}. \tag{3'}$$

In the algebraic sheet S_ν,

$$d\sigma = \frac{|\,dw\,|}{|\,w - A_\nu\,| + c_\nu}, \tag{3''}$$

where c_ν is defined by (1) and (2).

4. **Geodesic lines in the logarithmic ends.** We shall be interested here only in that family of geodesic lines which lie in the logarithmic ends and pass through $\tilde{A}_1$, the branch point over A_1. Let $\bar{P}_0^{(\nu)}$ be the point of Q_ν lying over P_0 of the w-plane. Let the polar coördinates of P_0 be (r_0, θ_0) $(-\pi < \theta_0 \leqq \pi)$, where the pole is taken at $w = 0$. To determine the geodesic line in the logarithmic ends through $\tilde{A}_1$ and $\bar{P}_0^{(\nu)}$, that curve C on W must be found which lies in the logarithmic ends, joins $\tilde{A}_1$ and $\bar{P}_0^{(\nu)}$, and for which the integral

$$\int_C \frac{|\,dw\,|}{|\,w\,|}$$

extended from $\tilde{A}_1$ to $\bar{P}_0^{(\nu)}$ is least. To this end consider the transformation of the logarithmic sheets by

$$\omega = \alpha + i\beta = \log w.$$

This maps the logarithmic ends on the ω-plane cut along the segment $(\log A_1, +\infty)$. This cut is the image of the junction lines of the logarithmic ends with the algebraic sheets. The sheet Q_ν is mapped on the strip $(2\nu - 1)\pi < \beta \leqq (2\nu + 1)\pi$ $(\nu = 0, \pm 1, \pm 2, \cdots)$.

Let C be any path in the logarithmic ends joining $\tilde{A}_1$ and $\bar{P}_0^{(\nu)}$, and let γ be the image of C in the ω-plane. γ is a curve joining the points with Cartesian coördinates $(\log A_1, 0)$, $(\log r_0, \theta_0 + 2\pi\nu)$. Since C stays in the logarithmic ends, γ does not cross the segment $(\log A_1, +\infty)$. The Euclidean length of γ is given by the integral of $|\,d\omega\,|$ along γ. But

$$\int_\gamma |\,d\omega\,| = \int_C \frac{|\,dw\,|}{|\,w\,|}$$

extended from $\tilde{A}_1$ to $\bar{P}_0^{(\nu)}$. Consequently the value of this latter integral will be least if and only if C is the image of the straight line of the ω-plane joining the points $(\log A_1, 0)$, $(\log r_0, \theta_0 + 2\pi\nu)$. This line is given parametrically, with parameter t, by

$$\omega = \omega(t) = \left(\log A_1 + t \log \frac{r_0}{A_1}\right) + it(\theta_0 + 2\pi\nu).$$

The image E_ν in the w-plane is given parametrically by

$$\log r = \log A_1 + t \log \frac{r_0}{A_1}, \qquad \theta = t(\theta_0 + 2\pi\nu).$$

If $\nu \neq 0$, $\theta_0 \neq 0$, this image will be the logarithmic spiral

$$r = A_1 \exp\left[\frac{\theta}{\theta_0 + 2\pi\nu} \log \frac{r_0}{A_1}\right],$$

which becomes the circle $r = A_1$ if $r_0 = A_1$. In the special case $\nu = 0$, $\theta_0 = 0$, the image will be the line $\theta = 0$.

From the above discussion we have the following result:[4]

The family of geodesic lines through the branch point over A_1 and lying in the logarithmic ends is the family of curves $\hat{E}_\nu$ on W lying over the spirals E_ν.

5. **Geodesic lines in an algebraic sheet.** We shall be interested here in those geodesic lines which lie in the algebraic sheet S_ν and join two points of S_ν lying on the same half-ray emanating from $\tilde{A}_\nu$, the branch point over A_ν. If $w - A_\nu = re^{i\theta}$, then

$$\int \frac{|\,dw\,|}{|\,w - A_\nu\,| + c_\nu} = \int \frac{(dr^2 + r^2 d\theta^2)^{\frac{1}{2}}}{r + c_\nu}. \tag{4}$$

If (r_1, θ_0), (r_2, θ_0) are two points of S_ν on the same half-ray emanating from $\tilde{A}_\nu$, and C is any path in S_ν joining these points

$$\int_C \frac{(dr^2 + r^2 d\theta^2)^{\frac{1}{2}}}{r + c_\nu} \geqq \int_C \frac{|\,dr\,|}{r + c_\nu} \geqq \left| \int_C \frac{dr}{r + c_\nu} \right|.$$

But the integral on the right is independent of the path and has the value $|\log (r_2 + c_\nu)/(r_1 + c_\nu)\,|$, which is the value of (4) taken along $\theta = \theta_0$ from (r_1, θ_0) to (r_2, θ_0). Hence we have the following:

The geodesic line in S_ν which joins two points of S_ν lying on the same half-ray emanating from the branch point over A_ν is the straight line joining these points.

The above conclusions are true if in particular $r_1 = 0$. We then have:

The family of geodesic lines in S_ν through the branch point over A_ν is the family of straight lines through that point.[5]

[4] If $w = re^{i\theta}$, $\int |\,dw\,| / |\,w\,|$ becomes $\int r^{-1}(dr^2 + r^2 d\theta^2)^{\frac{1}{2}}$. The extremals of this integral are found to be the spirals of the family $r = be^{a\theta}$. The spiral of this family determined by the sets of coördinates $(A_1, 0)$, $(r_0, \theta_0 + 2\pi\nu)$ is precisely the spiral E_ν found above.

[5] The extremals of integral (4) through $r = 0$ are found to be the lines of the family $\theta =$ constant.

6. **Distances in the various sheets.** The distance from $\tilde{A}_1$, the branch point over A_1, to $\bar{P}_0^{(\nu)}$, a point of the logarithmic sheet Q_ν, is given by

$$D(\tilde{A}_1, \bar{P}_0^{(\nu)}) = \int \frac{(dr^2 + r^2 d\theta^2)^{\frac{1}{2}}}{r},$$

where the integral is extended from $\tilde{A}_1$ to $\bar{P}_0^{(\nu)}$ along the curve $\bar{E}_\nu$. If $\bar{P}_0^{(\nu)}$ lies over the point $r_0 e^{i\theta_0}$ $(-\pi < \theta_0 \leqq \pi)$, along $\bar{E}_\nu$

$$r = A_1 \exp\left[\frac{\theta_0}{\theta_0'} \log \frac{r_0}{A_1}\right],$$

where $\theta_0' = \theta_0 + 2\pi\nu$, provided $\theta_0' \neq 0$. Evaluation of the integral expressing $D(\tilde{A}_1, \bar{P}_0^{(\nu)})$ along this curve yields

$$D(\tilde{A}_1, \bar{P}_0^{(\nu)}) = \left| \log \frac{w_0}{A_1} \right|, \tag{5}$$

where $w_0 = r_0 e^{i\theta_0'}$. In the special case $\theta_0' = 0$, $\bar{E}_\nu$ lies over the line $\theta = 0$, and it is immediate that (5) is still correct.

The distance from $\tilde{A}_\nu$, the branch point over A_ν, to a point $\bar{P}_\nu$ of the algebraic sheet S_ν is given by

$$D(\tilde{A}_\nu, \bar{P}_\nu) = \int \frac{(dr^2 + r^2 d\theta^2)^{\frac{1}{2}}}{r + c_\nu},$$

where the integral is extended along the straight line joining $\tilde{A}_\nu$ and $\bar{P}_\nu$, and the pole of the system is at $\tilde{A}_\nu$. If (r_0, θ_0) are the coördinates of $\bar{P}_\nu$,

$$D(\tilde{A}_\nu, \bar{P}_\nu) = \int_0^{r_0} \frac{dr}{r + c_\nu} = \log\left(1 + \frac{r_0}{c_\nu}\right). \tag{6}$$

The junction lines between the logarithmic sheet Q_0 and the algebraic sheet S_1, as well as the junction lines between successive algebraic sheets, are the loci of points which separate two regions in which there are different analytic expressions for the distance from the branch point from which the junction line emanates. For instance, if P_0, a point of either of the junction lines between Q_0 and S_1, is considered as lying on a shore of Q_0, the distance D from $\tilde{A}_1$ is given by (5), while if considered as lying on a shore of S_1, that distance is given by (6) with $\nu = 1$. In the first case

$$D = \log \frac{\rho_0}{A_1},$$

where ρ_0 is the radius vector of P_0 with the pole over $w = 0$. If we use (6) to determine the distance D' in the second case, r_0 is the radius vector of P_0 with the pole at $\tilde{A}_1$:

$$r_0 = \rho_0 - A_1,$$

and we have

$$D' = \log\left(1 + \frac{\rho_0 - A_1}{c_1}\right).$$

But by the notation of (1) and (2), $c_1 = p_0 = A_1$, and we again have

$$D' = \log\frac{\rho_0}{A_1} = D.$$

If P_0, a point of either of the junction lines between S_ν and $S_{\nu-1}$, is considered as lying on the shore of S_ν, the distance D from $\tilde{A}_\nu$ to P_0 is given by (6), where the radius vector of P_0 is measured from the pole at $\tilde{A}_\nu$:

$$D = \log\left(1 + \frac{r_0}{c_\nu}\right).$$

If P_0 is considered as lying on the shore of $S_{\nu-1}$, the distance D' from $\tilde{A}_\nu$ to P_0 is measured in the metric of $S_{\nu-1}$. But in this case $\tilde{A}_\nu$ and P_0 lie on the same half-ray emanating from the pole at $A_{\nu-1}$. Therefore by the first of the two results stated at the end of the preceding section

$$D' = \int_{r_1}^{r_2} \frac{(dr^2 + r^2 d\theta^2)^{\frac{1}{2}}}{r + c_\nu}$$

extended along $\theta =$ constant, where $r_1 = |A_\nu - A_{\nu-1}| = p_{\nu-1}$ and $r_2 = r_0 + p_{\nu-1}$. Hence, on recalling the notation given in (1) and (2), we see that

$$D' = \log\frac{r_0 + p_{\nu-1} + c_{\nu-1}}{p_{\nu-1} + c_{\nu-1}} = \log\left(1 + \frac{r_0}{c_\nu}\right) = D.$$

Similar considerations show that the distance from $\tilde{A}_{\nu+1}$ to a point of either of the junction lines between S_ν and $S_{\nu+1}$ is the same whether measured in the metric of S_ν or of $S_{\nu+1}$.

In other words, *the value of the distance from an algebraic branch point to a point of either of the junction lines emanating from that branch point is the same when measured in the metric of either of the sheets joined along that junction line.*

7. **Geodesic lines on the surface.** The problem of determining the shortest path on the surface from $\tilde{A}_1$, the branch point over A_1, to any point P of the surface now arises. Let G be the path on W defined as follows: If P is in the logarithmic sheet Q_ν, G is the spiral $\tilde{E}_\nu$ joining $\tilde{A}_1$ and P. If P is in the algebraic sheet S_1, G is the geodesic line in S_1 through $\tilde{A}_1$ and P. If P is in the algebraic sheet S_{ν_0} ($\nu_0 \geqq 2$), G is the path obtained by going from $\tilde{A}_1$ along the geodesic line in S_1 to the branch point over A_2, from this point along the geodesic line in S_2 to the branch point over A_3, and so on, coming finally to the branch point over A_{ν_0}. Then go from this point along the geodesic line in S_{ν_0} to P.

Now let C be any curve on the surface joining $\tilde{A}_1$ and P, along which the left

member of (7) (see below) exists. Let the portion of C in any sheet lying over a finite region of the w-plane be rectifiable. It will be shown that

$$\int_C \frac{|dw|}{\Delta(\tilde{w})} \geqq \int_G \frac{|dw|}{\Delta(\tilde{w})}. \tag{7}$$

Since C, with its end points, forms a closed set of points on W, it must lie in a finite number of sheets. In the first place, it is necessary to establish (7) only for those curves C such that the portion of C in any sheet lies over a bounded region of the w-plane, say over $|w| < R$. For, the point at infinity is not accessible along any path of finite length in the metric considered. From (5) it is clear that in the logarithmic sheets such distances become infinite like log $|w|$. Along a path C in the algebraic sheet S joining (r_0, θ_0), (r, θ)

$$\int_C d\sigma = \int_C \frac{(dr^2 + r^2 d\theta^2)^{\frac{1}{2}}}{r + c_\nu} \geqq \int_{r_0}^{r} \frac{|dr|}{r + c_\nu} \geqq \left| \int_{r_0}^{r} \frac{dr}{r + c_\nu} \right|.$$

This last integral is independent of the path and with increasing r becomes infinite like log r.

Moreover, it is necessary to consider only curves C such that there are in an algebraic sheet S_ν but a finite number of segments of C which join a point of the junction line emanating from the branch point over A_ν with a point of the junction line emanating from the branch point over $A_{\nu+1}$. For, the integral of $d\sigma$ along any curve which has an infinite number of such segments is not bounded. On one such segment Γ,

$$\int_\Gamma d\sigma \geqq \int_\Gamma \frac{|dw|}{R + |A_\nu| + c_\nu} = \frac{l}{R + |A_\nu| + c_\nu},$$

where l is the Euclidean length of Γ. But for each such segment we have $l = |A_\nu - A_{\nu-1}|$, and so if there are an infinite number of such segments, the above statement follows at once.

Suppose C crosses a junction line emanating from the algebraic branch point over A_ν an infinite number of times. All intersections of C with this junction line fall in a finite interval along the junction line. Moreover, there will be on the junction line a finite number of intervals I_k $(k = 1, \cdots, m)$ with the following properties: The endpoints of each I_k are intersections of C with the junction line. In each I_k there are an infinite number of intersections of C with the junction line, and the portion of C between the endpoints of I_k lies entirely in the two sheets which are joined along this junction line. Exterior to the intervals I_k there will be a finite number of intersections of C with the junction line.

Let α_k be the first point of I_k reached in traveling along C from $\tilde{A}_1$ to P, and β_k the second. Then, since the segment I_k is the geodesic line through α_k and β_k for either of the sheets joined along this junction line,

$$\int_{\alpha_k}^{\beta_k} d\sigma \geqq \int_{I_k} d\sigma, \tag{8}$$

where the integral on the left is extended along the portion of C between α_k and β_k.

Let C' be the curve on W obtained from C by replacing the portion of C between each pair of points α_k, β_k on all junction lines by the corresponding interval I_k. From (8) it follows that

$$\int_C d\sigma \geqq \int_{C'} d\sigma.$$

From this it follows that it is necessary to establish the validity of (7) only for such curves C which may fall along certain junction lines throughout a finite number of finite intervals but otherwise meet junction lines emanating from algebraic branch points in but a finite number of points.

Let C be such a curve joining $\bar{A}_1$ and P. Let $\{T_i\}$ $(i = 1, \cdots, n)$ be the sequence of regions through which C passes, where a region T_i is either a single algebraic sheet, or a sequence of logarithmic sheets such that in passing through the sequence along C, no two of the logarithmic sheets are separated by an algebraic sheet. Such a sequence of logarithmic sheets constituting a region T_i, $i < n$, must have Q_0 as the first and last sheet. It may in particular reduce to the single sheet Q_0. T_1 must be either the sheet S_1 or a sequence of logarithmic sheets. If P lies in the algebraic sheet S_{ν_0}, T_n is the sheet S_{ν_0}. If P lies in the logarithmic sheet Q_{ν_0}, T_n is a sequence of logarithmic sheets, the first of which is Q_0 and the last Q_{ν_0}. In forming the sequence $\{T_i\}$, if C coincides with a junction line emanating from an algebraic branch point throughout an interval, that portion of C can be considered as lying in either of the two sheets which are joined along that junction line.

For $1 < i \leqq n$, let $B_{i-1,i}$ be the point at which C enters T_i; then $B_{i,i+1}$, $i < n$, will be the point at which C leaves T_i and enters T_{i+1}. Let $B_{0,1}$ be $\bar{A}_1$. If $B_{i-1,i}$ lies on the junction line emanating from the branch point over A_ν, let $G_{i-1,i}$ be the geodesic line through that branch point and $B_{i-1,i}$. This geodesic line will fall along the junction line, if considered in the metric of either T_i or T_{i-1}. Since $B_{0,1}$ coincides with $\bar{A}_1$, the integral of $d\sigma$ along $G_{0,1}$ is 0. If T_i is the algebraic sheet S_ν, let G_i be the geodesic line in T_i through the branch point over A_ν and the branch point over $A_{\nu+1}$. Let C_i be the portion of C in T_i.

We shall now consider the integral of $d\sigma$ along C_i. In this connection it is convenient to distinguish several cases.

Case I. *Those regions T_i for which $B_{i-1,i}$ and $B_{i,i+1}$ lie on the same junction line.* For all such T_i, $i < n$. Any T_i consisting of a sequence of logarithmic sheets must come under this case. If $B_{i-1,i}$ lies between the branch point and $B_{i,i+1}$, or $B_{i-1,i}$ coincides with $B_{i,i+1}$,

$$\int_{G_{i-1,i}} d\sigma \leqq \int_{G_{i,i+1}} d\sigma$$

and

$$(9) \qquad \int_{C_i} d\sigma \geqq \int_{G_{i,i+1}} d\sigma - \int_{G_{i-1,i}} d\sigma;$$

for, the geodesic line through $B_{i-1,i}$ and $B_{i,i+1}$ in the metric of T_i falls along the junction line.

On the other hand, if

$$\int_{G_{i-1,i}} d\sigma > \int_{G_{i,i+1}} d\sigma,$$

(9) is still valid, for in that case the right member is negative. Hence for all regions T_i considered under Case I, (9) obtains.

Case II. *Those regions T_i for which $B_{i-1,i}$ and $B_{i,i+1}$ lie on different junction lines.* Here again for each such region T_i, $i < n$. Also, each region T_i of this class is an algebraic sheet S_ν. To treat this case it is necessary to recognize two cases according as C enters S_ν over a junction line emanating from the branch point over A_ν, or a junction line emanating from the branch point over $A_{\nu+1}$.

In the first case, since the straight line through $\bar{A}_\nu$ and $B_{i,i+1}$ is the geodesic line through these points,

$$\int_{G_{i-1,i}} d\sigma + \int_{C_i} d\sigma \geqq \int_{G_i} d\sigma + \int_{G_{i,i+1}} d\sigma,$$

or

$$(10) \qquad \int_{C_i} d\sigma \geqq \int_{G_i} d\sigma + \int_{G_{i,i+1}} d\sigma - \int_{G_{i-1,i}} d\sigma.$$

In the second case, since $G_{i,i+1}$ is the geodesic line through $B_{i,i+1}$ and the branch point over A_ν, we have

$$\int_{G_i} d\sigma + \int_{G_{i-1,i}} d\sigma + \int_{C_i} d\sigma \geqq \int_{G_{i,i+1}} d\sigma,$$

or

$$(11) \qquad \int_{C_i} d\sigma \geqq - \int_{G_i} d\sigma + \int_{G_{i,i+1}} d\sigma - \int_{G_{i-1,i}} d\sigma.$$

There still remains for consideration

Case III. *The region T_n.* If P is in the logarithmic sheet Q_ν, let g_n be the geodesic line $\bar{E}_\nu$ in the logarithmic ends through $\bar{A}_1$ and P. If P is in the algebraic sheet S_ν, let g_n be the geodesic line in S_ν through P and the branch point over A_ν. On considering the integral of $d\sigma$ along C_n, we shall first treat the case in which T_n is a sequence of logarithmic sheets or the algebraic sheet

S_ν and C enters T_n over a junction line emanating from the branch point over A_ν. In this case, we have

$$\int_{G_{n-1,n}} d\sigma + \int_{C_n} d\sigma \geqq \int_{g_n} d\sigma,$$

or

$$\int_{C_n} d\sigma \geqq \int_{g_n} d\sigma - \int_{G_{n-1,n}} d\sigma. \tag{12}$$

If T_n is the algebraic sheet S_ν and C enters T_n over a junction line emanating from the branch point over $A_{\nu+1}$, since g_n is the geodesic line in T_n through P and the branch point over A_ν, we have

$$\int_{G_{n-1,n}} d\sigma + \int_{G_n} d\sigma + \int_{C_n} d\sigma \geqq \int_{g_n} d\sigma,$$

or

$$\int_{C_n} d\sigma \geqq - \int_{G_{n-1,n}} d\sigma - \int_{G_n} d\sigma + \int_{g_n} d\sigma. \tag{13}$$

We then have that for each $1 \leqq i \leqq n$, the integral $\int_{C_i} d\sigma$ satisfies one of the five inequalities (9), (10), (11), (12), (13).

In the following discussion, if any region T_i is an algebraic sheet S_ν and C leaves this region over a junction line different from the one over which it entered the region, we shall say that C *crosses* the region T_i.

In the first place we have that

$$\int_C d\sigma = \sum_{i=1}^{n-1} \int_{C_i} d\sigma + \int_{C_n} d\sigma. \tag{14}$$

We shall now replace each term of the summation on the right by the right member of the one of the three inequalities (9), (10) or (11) which is valid for the T_i in question. The result is

$$\sum_{i=1}^{n-1} \int_{C_i} d\sigma \geqq \sum_{i=1}^{n-1}{}' \left(\pm \int_{G_i} d\sigma \right) + \sum_{i=1}^{n-1} \left(\int_{G_{i,i+1}} d\sigma - \int_{G_{i-1,i}} d\sigma \right), \tag{15}$$

where $\sum'$ denotes the deleted sum in which a term appears only for those values of i for which either (10) or (11) is in force. Moreover, the sign of the term is positive when (10) is in force and negative when (11) is.

From the result stated at the end of §6, it follows that

$\int_{G_{i,i+1}} d\sigma$ for any index i_0 is equal to $\int_{G_{i-1,i}} d\sigma$ for the next greater index $i_0 + 1$.

Therefore, since the integral of $d\sigma$ along $G_{0,1}$ is zero,

$$\sum_{i=1}^{n-1}\left(\int_{G_{i,i+1}} d\sigma - \int_{G_{i-1,i}} d\sigma\right) = \int_{G_{n-1,n}} d\sigma.$$

If this is used in (15), (14) becomes

$$\int_C d\sigma \geqq \sum_{i=1}^{n-1}{}'\left(\pm \int_{G_i} d\sigma\right) + \int_{G_{n-1,n}} d\sigma + \int_{C_n} d\sigma. \tag{16}$$

It is again convenient to distinguish several cases.

Case I. *P lies in a logarithmic sheet.* Then T_n is a sequence of logarithmic sheets. In that case (12) is satisfied. Moreover, if P lies in the logarithmic sheet Q_ν, g_n is the geodesic line $\tilde{E}_\nu$, and (16) becomes

$$\int_C d\sigma \geqq \sum_{i=1}^{n-1}{}'\left(\pm \int_{G_i} d\sigma\right) + \int_{\tilde{E}_\nu} d\sigma. \tag{16$_1$}$$

It should first be noticed that, for the case under consideration, if C crosses any algebraic sheet S_ν, contributing a term to the deleted sum, it crosses S_ν an even number of times, so that S_ν appears an even number of times in the sequence $\{T_i\}$. For half of the regions $T_{i'}$ which are the sheet S_ν, C enters $T_{i'}$ from the logarithmic sheet Q_0 or an algebraic sheet of lower index, while for the other half, from an algebraic sheet of higher index. Consequently, for half of these regions (10) is in force while for the other half (11) is. Hence, corresponding to each term of the deleted sum which carries a negative sign, there is a term numerically equal with a positive sign, and

$$\sum_{i=1}^{n-1}{}'\left(\pm \int_{G_i} d\sigma\right) = 0.$$

(16$_1$) then becomes

$$\int_C d\sigma \geqq \int_{\tilde{E}_\nu} d\sigma.$$

But for this case $\tilde{E}_\nu$ is precisely the path G along which the integral on the right side of (7) is extended, and hence (7) has been established if P is in a logarithmic sheet.

Case II. *P lies in an algebraic sheet S_{ν_0} and C enters S_{ν_0} from S_{ν_0-1}.*[6] In this case $\int_{C_n} d\sigma$ again satisfies (12). If we let g_{ν_0} be the geodesic line in S_{ν_0} through P and the branch point over A_{ν_0}, g_{ν_0} is identical with g_n, and (16) becomes

$$\int_C d\sigma \geqq \sum_{i=1}^{n-1}{}'\left(\pm \int_{G_i} d\sigma\right) + \int_{g_{\nu_0}} d\sigma. \tag{16$_2$}$$

[6] If $\nu_0 = 1$, C enters S_1 from Q_0.

For the case under consideration C must cross each algebraic sheet S_ν $(1 \leqq \nu < \nu_0)$ an odd number of times and an algebraic sheet S_ν $(\nu \geqq \nu_0)$ an even number of times. Consequently a sheet S_ν $(1 \leqq \nu < \nu_0)$ appears in the sequence $\{T_i\}$ an odd number of times, say $2m + 1$. In m of these regions which are the sheet S_ν, C enters S_ν from $S_{\nu+1}$ so that (11) is in force, while in the remaining $m + 1$, C enters S_ν from $S_{\nu-1}$, or Q_0 if $\nu_0 = 1$, so that (10) is in force. Hence, if we consider the terms of the deleted sum for which the region T_i is an algebraic sheet S_ν $(1 \leqq \nu < \nu_0)$ corresponding to each term with a negative sign, there will be a term numerically equal with a positive sign. In addition there will be just one remaining positive term. Consequently the terms of the deleted sum for which the T_i are algebraic sheets S_ν $(1 \leqq \nu < \nu_0)$ add up to

$$\sum_{\nu=1}^{\nu_0-1} \int_{G_\nu} d\sigma,$$

where G_ν is the geodesic line in S_ν through the branch point over A_ν and the branch point over $A_{\nu+1}$.

For the terms of the deleted sum for which T_i is an algebraic sheet S_ν $(\nu \geqq \nu_0)$, since each such S_ν is crossed by C an even number of times, the discussion is exactly the same as that of the deleted sum in Case I. Therefore the sum of those terms is zero, and we have

$$\sum_{i=1}^{n-1}{}' \left(\pm \int_{G_i} d\sigma \right) = \sum_{\nu=1}^{\nu_0-1} \int_{G_\nu} d\sigma.$$

If this is used in (16_2), we have

$$\int_C d\sigma \geqq \sum_{\nu=1}^{\nu_0-1} \int_{G_\nu} d\sigma + \int_{g_{\nu_0}} d\sigma.$$

But when P lies in S_{ν_0} the path made up of $G_1, G_2, \cdots, G_{\nu_0-1}, g_{\nu_0}$ is precisely the path G along which the integral in the right member of (7) is extended, and with this (7) has been proved in the present case.

There remains for consideration

Case III. *P lies in S_{ν_0} and C enters S_{ν_0} from S_{ν_0+1}.* Here (13) is satisfied. In the notation of the preceding case, g_n is g_{ν_0}, and G_n is G_{ν_0}. (16) then becomes

$$(16_3) \qquad \int_C d\sigma \geqq \sum_{i=1}^{n-1}{}' \left(\pm \int_{G_i} d\sigma \right) - \int_{G_{\nu_0}} d\sigma + \int_{g_{\nu_0}} d\sigma.$$

In this case C crosses each algebraic sheet S_ν $(1 \leqq \nu \leqq \nu_0)$ an odd number of times and an algebraic sheet S_ν $(\nu > \nu_0)$ an even number of times. The discussion of the deleted sum is exactly the same as that of Case II, with the modification that here the sum of the integrals extends from $\nu = 1$ to $\nu = \nu_0$ instead of from $\nu = 1$ to $\nu = \nu_0 - 1$ as in Case II. Consequently (16_3) becomes

$$\int_C d\sigma \geqq \sum_{\nu=1}^{\nu_0} \int_{G_\nu} d\sigma - \int_{G_{\nu_0}} d\sigma + \int_{g_{\nu_0}} d\sigma = \sum_{\nu=1}^{\nu_0-1} \int_{G_\nu} d\sigma + \int_{g_{\nu_0}} d\sigma.$$

But again this is precisely inequality (7), and with this the proof of (7) is complete in all cases.

The proof of (7) establishes the fact that *the path G through $\bar{A}_1$, the branch point over A_1, and any point P of the surface is the geodesic line on the surface through $\bar{A}_1$ and P.*

8. **Distance from the branch point over A_1 to the branch point over A_{ν_0}.** From the result of the preceding section, it follows that the distance from the branch point over A_1 to the branch point over A_{ν_0} is given by

$$D(\bar{A}_1, \bar{A}_{\nu_0}) = \sum_{\nu=1}^{\nu_0-1} \int_{G_\nu} d\sigma.$$

But in the notation given by (1) and (2),

$$\int_{G_\nu} d\sigma = \int_0^{p_\nu} \frac{dr}{r + c_\nu} = \log\left(1 + \frac{p_\nu}{c_\nu}\right) = \log \frac{c_{\nu+1}}{c_\nu},$$

and we have

$$D(\bar{A}_1, \bar{A}_{\nu_0}) = \sum_{\nu=1}^{\nu_0-1} \log \frac{c_{\nu+1}}{c_\nu} = \log \frac{c_{\nu_0}}{p_0}. \tag{17}$$

9. **Lengths of non-Euclidean circles in the various sheets.** Consider the locus K_0, in the logarithmic ends, of all points P of W which are at the non-Euclidean distance ρ from the branch point over A_1. If P is given by $w = re^{i\theta}$, from (5) this locus is given by $|\log (w/A_1)| = \rho$. This curve on W lies over the curve

$$r = A_1 \exp\left[\pm(\rho^2 - \theta^2)^{\frac{1}{2}}\right]$$

of the w-plane. If K_0' is the branch of this curve given by the positive sign and K_0'' that by the negative, the non-Euclidean length of K_0, is given by

$$L_0(\rho) = \int_{K_0'} d\sigma + \int_{K_0''} d\sigma = 2\rho \int_{-\rho}^{\rho} \frac{d\theta}{(\rho^2 - \theta^2)^{\frac{1}{2}}};$$

that is,

$$L_0(\rho) = 2\pi\rho. \tag{18}$$

Now consider the locus K_ν, in the algebraic sheet S_ν, of points P which are at a non-Euclidean distance ρ from the branch point over A_ν. If (r, θ) are the coördinates of P, where the pole is taken at the branch point over A_ν, it is found from (6) that this locus is given by $r = c_\nu(e^\rho - 1)$. Let $L_\nu(\rho)$ be the non-Euclidean length of K_ν. Then

$$L_\nu(\rho) = (1 - e^{-\rho}) \int_0^{2\pi} d\theta = 2\pi(1 - e^{-\rho}). \tag{19}$$

10. **The regions W_ρ on the surface and the function $L(\rho)$.** Let W_ρ be that region of the surface W which consists of those points of W whose N.E.-distance from the branch point over A_1 does not exceed the positive number ρ. Let Γ_ρ be the boundary of W_ρ. Γ_ρ is obtained by measuring out a distance ρ along all geodesic lines on W through $\tilde{A}_1$. If in this process a branch point is encountered, the measuring must be continued out along all geodesic lines through that branch point in all sheets intertwined at that point until the complete distance ρ is attained. It follows from (17) and the result of §7 that if $\log (c_{\nu_0}/p_0) < \rho < \log (c_{\nu_0+1}/p_0)$, Γ_ρ consists of $\nu_0 + 1$ N.E.-circles. First there is K_0 in the logarithmic ends with N.E.-radius ρ; then for each $1 \leqq \nu \leqq \nu_0$, there is K_ν in S_ν with N.E.-radius $\rho - \log (c_\nu/p_0)$. Therefore, if $L(\rho)$ is the length of the boundary of W_ρ, from (18) and (19)

$$L(\rho) = 2\pi\rho + \sum_{\nu=1}^{\nu_0} 2\pi\left[1 - \exp\left(-\rho + \log\frac{c_\nu}{p_0}\right)\right]. \tag{20}$$

It is convenient to represent the summation in (20) by a Stieltjes integral. Let $n(t)$ be the number of algebraic branch points of the surface at a N.E.-distance less than t from $\tilde{A}_1$. $n(t)$ is a step-function with discontinuities at the points $t = \log (c_\nu/p_0)$ ($\nu = 1, 2, 3, \cdots$). At each of these points the saltus is $+1$:

$$n(t) = \nu, \qquad \log (c_\nu/p_0) < t \leqq \log (c_{\nu+1}/p_0) \qquad (\nu = 1, 2, 3, \cdots).$$

The Stieltjes integral

$$2\pi \int_0^\rho (1 - e^{-\rho+t})\, dn(t)$$

exists and is equal to the summation appearing in (20), where $\log (c_{\nu_0}/p_0) < \rho \leqq \log (c_{\nu_0+1}/p_0)$.

The expression for $L(\rho)$ then becomes

$$L(\rho) = 2\pi\left[\rho + \int_0^\rho (1 - e^{-\rho+t})\, dn(t)\right],$$

and after an integration by parts

$$L(\rho) = 2\pi\left[\rho + \int_0^\rho n(t)e^{-\rho+t}\, dt\right]. \tag{21}$$

11. **Sufficient conditions that the surface be parabolic.** In order to obtain sufficient conditions that W be parabolic, the following lemma will be useful:

LEMMA I. *If $f(x)$ and $g(x)$ are two positive monotone increasing functions of the real variable x defined for all $x > 0$, and if $g(x) = O(x)$ as x becomes infinite, the integrals*

$$\int_a^\infty \frac{dx}{f(x) + g(x)} \quad \text{and} \quad \int_a^\infty \frac{dx}{f(x)}, \qquad a > 0,$$

diverge together.

It is at once evident that the second of these integrals diverges if the first does. In order to prove that the first diverges whenever the second does, we first suppose there exist positive constants x_0 and m such that $f(x)/g(x) > m$ for $x > x_0$. Then

$$\frac{1}{f(x) + g(x)} > \frac{m}{m+1} \cdot \frac{1}{f(x)}$$

for $x > x_0$, and from this the result follows.

If this condition on $f(x)/g(x)$ is not met, there exists an infinite sequence $\{x_n\}$, where $x_n \to \infty$ as $n \to \infty$, such that $f(x_n)/g(x_n) < n^{-1}$ for each positive integer n. But then, due to the monotone character of $f(x)$ and $g(x)$, it follows that

$$\int_{\frac{1}{2}x_n}^{x_n} \frac{dx}{f(x) + g(x)} > \frac{n}{n+1} \cdot \frac{x_n}{2g(x_n)}. \tag{22}$$

Since $g(x) = O(x)$, the right member of (22) remains greater than some positive constant independent of n. But (22) then contradicts the Cauchy condition for the convergence of the first integral in the lemma. With this the proof of the lemma is complete.

A sufficient condition that W be parabolic can now be obtained. Since $n(t)$ is a monotone increasing function of t, from (21) it follows that

$$L(\rho) < 2\pi \left[\rho + n(\rho) \int_0^\rho e^{-\rho+t}\, dt \right] < 2\pi[\rho + n(\rho)]$$

and

$$\frac{1}{L(\rho)} > \frac{1}{2\pi[\rho + n(\rho)]}.$$

Hence the integral

$$\int^\infty \frac{d\rho}{L(\rho)} \text{ will diverge if } \int^\infty \frac{d\rho}{\rho + n(\rho)} \text{ diverges.}$$

By Lemma I we can then say that

$$\int^\infty \frac{d\rho}{L(\rho)} \text{ diverges if } \int^\infty \frac{d\rho}{n(\rho)} \text{ does.}$$

On the other hand, from (21) we also have

$$L(\rho) > 2\pi \left[\rho + n(\tfrac{1}{2}\rho) \int_{\frac{1}{2}\rho}^\rho e^{-\rho+t}\, dt \right] = 2\pi[\rho + n(\tfrac{1}{2}\rho)(1 - e^{-\frac{1}{2}\rho})].$$

Hence, for sufficiently large ρ,

$$\frac{1}{L(\rho)} < \frac{1}{\pi[2\rho + n(\frac{1}{2}\rho)]},$$

from which it follows that

$$\text{if } \int^{\infty} \frac{d\rho}{L(\rho)} \text{ diverges, so also does } \int^{\infty} \frac{d\rho}{4\rho + n(\rho)}.$$

From Lemma I it can then be concluded that

$$\int^{\infty} \frac{d\rho}{n(\rho)} \text{ diverges if } \int^{\infty} \frac{d\rho}{L(\rho)} \text{ does.}$$

From the criterion of Ahlfors stated in the introduction we now have

THEOREM I. *Let W be an open simply-connected Riemann surface of the class defined in* §2. *A sufficient condition that W be parabolic is the divergence of the integral*

$$\int^{\infty} \frac{d\rho}{n(\rho)}.$$

Since it has been proved that the divergence of the integral of Theorem I is entirely equivalent to the divergence of the above integral, *the result stated in Theorem* I *is the best which can be obtained from the metric considered.*

A test in terms of the divergence of a series will be useful. Let

$$\rho_\nu = \log (c_\nu/p_0) \qquad (\nu = 1, 2, 3, \cdots).$$

By (1) and (2) $c_1 = p_0$, so that $\rho_1 = 0$. Moreover, $n(\rho) = 1$, $0 < \rho \leqq \rho_2$. Therefore

$$\int_0^{\rho_{\nu_0}} \frac{d\rho}{n(\rho)} = \rho_2 + \int_{\rho_2}^{\rho_{\nu_0}} \frac{d\rho}{n(\rho)}.$$

After an integration of the integral on the right by parts, we have

$$\int_0^{\rho_{\nu_0}} \frac{d\rho}{n(\rho)} = \frac{\rho_{\nu_0}}{\nu_0 - 1} - \int_{\rho_2}^{\rho_{\nu_0}} \rho d\left[\frac{1}{n(\rho)}\right]. \tag{23}$$

The function $1/n(\rho)$ is a step-function with discontinuities at the points ρ_ν ($\nu = 1, 2, 3, \cdots$). The saltus at ρ_ν is $\nu^{-1} - (\nu - 1)^{-1}$. Consequently

$$\int_{\rho_2}^{\rho_{\nu_0}} \rho d\left[\frac{1}{n(\rho)}\right] = -\sum_{\nu=2}^{\nu_0-1} \frac{\rho_\nu}{\nu(\nu - 1)},$$

and (23) becomes

$$\int_0^{\rho_{\nu_0}} \frac{d\rho}{n(\rho)} = \sum_{\nu=2}^{\nu_0} \frac{\rho_\nu}{\nu(\nu - 1)} + \frac{\rho_{\nu_0}}{\nu_0}. \tag{24}$$

Since each ρ_ν ($\nu \geqq 2$) is positive, it is immediate from (24) that if the series

$$\sum_{\nu=2}^{\infty} \frac{\rho_\nu}{\nu(\nu - 1)}$$

diverges, the integral

$$\int^{\infty} \frac{d\rho}{n(\rho)}$$

will also.

On the other hand, if the integral diverges, the quantity on the right side of (24) must become infinite with increasing ν_0. But, since the summation appearing in (24) is a monotone increasing function of ν_0, it can be concluded that at least one of the terms on the right must become infinite.

Now, if $\rho_{\nu_0}/\nu_0 \to \infty$ as $\nu_0 \to \infty$, corresponding to any constant $M > 0$, there is a positive number $N > 2$ such that $\rho_{\nu_0}/\nu_0 > M$ for $\nu_0 > N$, and hence

$$\frac{\rho_{\nu_0}}{\nu_0(\nu_0 - 1)} > \frac{M}{\nu_0 - 1} \quad \text{for } \nu_0 > N.$$

But then the series diverges.

If ρ_{ν_0}/ν_0 does not become infinite with increasing ν_0, the sum of the series to ν_0 terms must; that is, the series diverges. So in any case if the integral diverges, the series does also.

With this it has been proved that the divergence of the series is equivalent to the divergence of the integral. But this series is comparable to the series

$$\sum_{\nu=1}^{\infty} \frac{\rho_\nu}{\nu^2}.$$

Hence we have

THEOREM II. *A sufficient condition that W be of parabolic type is the divergence of the series*

$$\sum_{\nu=1}^{\infty} \frac{1}{\nu^2} \log \frac{c_\nu}{p_0}.$$

From Theorem II there can be obtained some information concerning the asymptotic behavior of the quantities $|A_\nu|$ which will guarantee that W be parabolic. In the first place it will be shown that if $|A_\nu| = O(\nu^n)$, where n is a positive integer, the series of Theorem II converges. In that case there is a positive constant M such that $|A_\nu| < M\nu^n$ for all ν. Then

$$\begin{aligned} c_\nu &= p_0 + p_1 + \cdots + p_{\nu-1} \\ &= 2|A_1| + \cdots + 2|A_{\nu-1}| + |A_\nu| < 2M(1^n + 2^n + \cdots + \nu^n). \end{aligned}$$

But, the sum of the n-th powers of the first ν integers is a polynomial of degree $n + 1$ in ν, so that $c_\nu/p_0 = O(\nu^{n+1})$ and $\log (c_\nu/p_0) = O (\log \nu)$. From this it follows that the series of Theorem II is dominated by the series $\sum_{\nu=1}^{\infty} k\nu^{-2} \log \nu$, where k is a positive constant. This latter series converges and hence the series of Theorem II does also. In other words, in order to have the present conditions for the parabolic case met, the points A_ν over which the algebraic

branch points of the surface lie must spread out more rapidly than any power of ν.

On the other hand, it will now be shown that if

$$| A_\nu | \sim \frac{1}{\log(\nu+1)} \exp\left[\frac{\nu+1}{\log(\nu+1)}\right],$$

the series of Theorem II diverges. In the first place, it is clear that the derivative of $(x+1)/\log(x+1)$ is asymptotically equal to $1/\log(x+1)$. From this we have

$$| A_\nu | \sim D_\nu \left[\frac{\nu+1}{\log(\nu+1)}\right] \exp\left[\frac{\nu+1}{\log(\nu+1)}\right].$$

Therefore for any positive constant $m' < 1$, there exists an integer $\nu_0 > 1$ such that

$$| A_\nu | > m' D_\nu \left[\frac{\nu+1}{\log(\nu+1)}\right] \exp\left[\frac{\nu+1}{\log(\nu+1)}\right].$$

Then

$$\frac{c_\nu}{p_0} = \frac{1}{p_0}(2 | A_1 | + \cdots + 2 | A_{\nu-1} | + | A_\nu |)$$

$$> K + m \sum_{k=\nu_0+1}^{\nu-1} D_k \left[\frac{k+1}{\log(k+1)}\right] \exp\left[\frac{k+1}{\log(k+1)}\right],$$

where $K = (2/p_0)(| A_1 | + \cdots + | A_{\nu_0} |)$ and $m = 2m'/p_0$.

If $f(x)$ is a positive monotone increasing function defined for $x > 0$,

$$\sum_{k=\nu_0+1}^{\nu-1} f(k) \geqq \int_{\nu_0}^{\nu-1} f(x)\, dx.$$

Therefore

$$\frac{c_\nu}{p_0} > M + m \exp\left[\frac{\nu}{\log \nu}\right],$$

where

$$M = K - m \exp\left[\frac{\nu_0+1}{\log(\nu_0+1)}\right], \qquad \text{a constant.}$$

But then there is an integer $\nu_0' > \nu_0$ such that for $\nu > \nu_0'$

$$\frac{c_\nu}{p_0} > \frac{1}{2} m \exp\left[\frac{\nu}{\log \nu}\right].$$

Hence the series of Theorem II for $\nu > \nu_0'$ is term by term greater than the series

$$\sum_{\nu=2}^{\infty} \left(\frac{1}{\nu \log \nu} + \frac{1}{\nu^2} \log \frac{1}{2} m\right).$$

But this series is comparable to a divergent series. Therefore the series of Theorem II diverges and the surface is parabolic.

From the above discussion it is clear that in order to obtain this result it is only necessary that there exist a positive constant m and a positive integer ν_0 such that for $\nu > \nu_0$

$$|A_\nu| > \frac{m}{\log(\nu+1)} \exp\left[\frac{\nu+1}{\log(\nu+1)}\right].$$

This result is of the nature of a comparison test. The above facts will be stated as

THEOREM III. *In order that the conditions for the parabolic case expressed in Theorems* I *and* II *be met,* $|A_\nu|$ *must become infinite more rapidly than any power of* ν. *But, if there exist a positive constant* m *and a positive integer* ν_0 *such that for* $\nu > \nu_0$

$$|A_\nu| > \frac{m}{\log(\nu+1)} \exp\left[\frac{\nu+1}{\log(\nu+1)}\right],$$

the surface W *is of parabolic type.*[7]

Other series the divergence of which is equivalent to the divergence of the series of Theorem II can be obtained. Two such series are

$$\sum_{\nu=1}^{\infty} \frac{1}{\nu} \log(1 + p_\nu/c_\nu) \quad \text{and} \quad \sum_{\nu=1}^{\infty} [(1 + p_\nu/c_\nu)^{1/\nu} - 1].$$

12. **The Gamma surface.** Let W be the open simply-connected Riemann surface on which the finite z-plane is mapped by

$$w = w(z) = \frac{1}{\Gamma(z)},$$

where $\Gamma(z)$ is the Gamma function defined by the Weierstrass product

$$\frac{1}{\Gamma(z)} = ze^{\gamma z}\left[\prod_{n=1}^{\infty}\left(1 + \frac{z}{n}\right)e^{-z/n}\right],$$

γ being the Euler constant.

This surface has a logarithmic branch point over $w = 0$ along with an infinite number of smooth sheets, and a transcendental singularity over $w = \infty$ which is a limit point of algebraic branch points of first order. There are no other singularities.[8]

[7] See footnote 3.

[8] For a discussion of these facts concerning the general structure of this surface see A. A. Utzinger, *Die reellen Züge der Riemann'schen Zetafunktion*, Inauguraldissertation, Zürich (1934). In section III of this paper the surface defined by $w = \Gamma(z)$ is discussed, and the above facts follow at once.

The algebraic branch points of the surface lie over the points of the w-plane corresponding to the zeros of

$$w'(z) = -\frac{\Gamma'(z)}{\Gamma^2(z)} = -\frac{\psi(z)}{\Gamma(z)},$$

where $\psi(z) = D_z \log \Gamma(z)$. $w'(z)$ has a removable singularity at each of the poles of $\Gamma(z)$ and when properly defined is not zero at these points.[9] The zeros of $w'(z)$ coincide with the zeros of $\psi(z)$.

The zeros of $\psi(z)$ are all simple and lie on the real axis. There is a single positive zero: $a_1 = 1.4616 \cdots$. There is in each of the intervals $(-\nu + 1, -\nu + 2)$ $(\nu = 2, 3, 4, \cdots)$ one zero a_ν. We can write[10]

$$a_\nu = -\nu + 1 + h_\nu,$$

where

$$h_\nu = \frac{1 + \delta_\nu}{\log(\nu - 2)}, \quad \nu > 2, \qquad \text{and} \qquad \lim_{\nu\to\infty} \delta_\nu = 0.$$

If we use the Euler formula $\Gamma(z)\Gamma(1 - z) = \pi/\sin \pi z$, we get

$$w(a_\nu) = \frac{\sin(-\nu + 1 + h_\nu)\pi}{\pi}\Gamma(\nu - h_\nu) = \frac{(-1)^{\nu-1}}{\pi}\Gamma(\nu - h_\nu)\sin \pi h_\nu, \; \nu > 2.$$

Since

$$\lim_{\nu\to\infty} \frac{\sin \pi h_\nu}{\pi h_\nu} = 1,$$

$$| w(a_\nu) | \sim h_\nu \Gamma(\nu - h_\nu) \sim \frac{\Gamma(\nu - h_\nu)}{\log \nu}.$$

Therefore $\lim_{\nu\to\infty} | w(a_\nu) | = \infty$.

If we let $A_\nu = w(a_\nu)$, the points of the set $\{A_\nu\}$ lie on the real axis of the w-plane since the Gamma function is real for real argument. Moreover, $A_\nu > 0$ for ν odd and < 0 for ν even. The set $\{A_\nu\}$ constitutes the points of the w-plane over which the algebraic branch points of the surface lie. Since each of the algebraic branch points is of first order, the Gamma surface belongs to the class defined in §2. We already had above that

$$| A_\nu | \sim \frac{\Gamma(\nu - h_\nu)}{\log \nu}. \tag{25}$$

If m and M are any two constants such that $0 < m < 1 < M$, there exists an integer $\nu_1' > 4$ such that for $\nu > \nu_1'$

$$m\frac{\Gamma(\nu - h_\nu)}{\log \nu} < | A_\nu | < M\frac{\Gamma(\nu - h_\nu)}{\log \nu}.$$

[9] See, e.g., N. Nielsen, *Handbuch der Theorie der Gammafunktion*, 1906.

[10] This result is due to Hermite, *Journal für Mathematik*, vol. 90(1881), pp. 332–338.

Also, since the h_ν's tend to zero with increasing ν, there exists a positive integer ν_1'' such that for $\nu > \nu_1''$, $h_\nu < \frac{1}{2}$. Then if ν_1 is an integer greater than both ν_1' and ν_1'', for $\nu > \nu_1$

$$\text{(26)} \qquad \frac{2\,|\,A_1\,| + \cdots + 2\,|\,A_{\nu-1}\,|}{|\,A_\nu\,|} < \frac{K}{|\,A_\nu\,|} + \frac{2M}{m} \sum_{k=1}^{\nu-\nu_1-1} \frac{\log \nu}{\log (\nu_1 + k)} \cdot \frac{\Gamma(\nu_1 + k)}{\Gamma(\nu - \frac{1}{2})},$$

where $K = 2\,|\,A_1\,| + \cdots + 2\,|\,A_{\nu_1}\,|$, a constant. Now, for $\nu > 2\nu_1$

$$\sum_{k=1}^{\nu-\nu_1-1} \frac{\log \nu}{\log (\nu_1 + k)} \cdot \frac{\Gamma(\nu_1 + k)}{\Gamma(\nu - \frac{1}{2})} = \frac{\log \nu}{\log (\nu - 1)} \cdot \frac{\Gamma(\nu - 1)}{\Gamma(\nu - \frac{1}{2})} + \frac{\log \nu}{\log (\nu - 2)} \cdot \frac{\Gamma(\nu - 2)}{\Gamma(\nu - \frac{1}{2})}$$

$$+ \frac{\log \nu}{\log (\nu - 3)} \cdot \frac{\Gamma(\nu - 3)}{\Gamma(\nu - \frac{1}{2})} + \sum_{k=1}^{\nu-\nu_1-4} \frac{\log \nu}{\log (\nu_1 + k)} \cdot \frac{\Gamma(\nu_1 + k)}{\Gamma(\nu - \frac{1}{2})}.$$

$$\Gamma(\nu - \tfrac{1}{2}) = (\nu - \tfrac{3}{2})(\nu - \tfrac{5}{2})(\nu - \tfrac{7}{2})\Gamma(\nu - \tfrac{7}{2}),$$

and for each set of values satisfying $1 \leqq k \leqq \nu - \nu_1 - 4$

$$\frac{\Gamma(\nu_1 + k)}{\Gamma(\nu - \frac{7}{2})} < 1.$$

Therefore

$$\frac{\log \nu}{\log (\nu_1 + k)} \cdot \frac{\Gamma(\nu_1 + k)}{\Gamma(\nu - \frac{1}{2})} < \frac{\log \nu}{(\nu - \frac{3}{2})(\nu - \frac{5}{2})(\nu - \frac{7}{2})}.$$

Moreover, there is a positive integer ν_2 such that for $\nu > \nu_2$

$$\frac{\log \nu}{\log (\nu - 1)} < \frac{\log \nu}{\log (\nu - 2)} < \frac{\log \nu}{\log (\nu - 3)} < 2.$$

Hence, if ν_0 is an integer greater than both $2\nu_1$ and ν_2, for $\nu > \nu_0$

$$\sum_{k=1}^{\nu-\nu_1-1} \frac{\log \nu}{\log (\nu_1 + k)} \cdot \frac{\Gamma(\nu_1 + k)}{\Gamma(\nu - \frac{1}{2})} < \frac{6\Gamma(\nu - 1)}{\Gamma(\nu - \frac{1}{2})} + \frac{(\nu - \nu_1 - 4) \log \nu}{(\nu - \frac{3}{2})(\nu - \frac{5}{2})(\nu - \frac{7}{2})}.$$

But each term on the right tends toward zero as $\nu \to \infty$. Therefore, since $K/|\,A_\nu\,|$ also approaches zero with increasing ν, we have from (26) that

$$\lim_{\nu \to \infty} \frac{2\,|\,A_1\,| + \cdots + 2\,|\,A_{\nu-1}\,|}{|\,A_\nu\,|} = 0.$$

By definition, $c_\nu = 2\,|\,A_1\,| + \cdots + 2\,|\,A_{\nu-1}\,| + |\,A_\nu\,|$. Then from the limit above

$$\lim_{\nu \to \infty} \frac{c_\nu}{|\,A_\nu\,|} = 1$$

and

$$c_\nu \sim \frac{\Gamma(\nu - h_\nu)}{\log \nu}.$$

But then

$$\log (c_\nu / p_0) \sim \log \Gamma(\nu - h_\nu) - \log \log \nu.$$

By the Stirling formula,[11]

$$\log \Gamma(\nu - h_\nu) = (\nu - h_\nu - \tfrac{1}{2}) \log (\nu - h_\nu) - \nu + h_\nu + \log (2\pi)^{\frac{1}{2}} + \frac{\theta}{12(\nu - h_\nu)},$$

where $0 \leqq \theta < 1$. From this it follows that

$$\log (c_\nu/p_0) \sim \nu \log \nu. \tag{27}$$

The function $n(\rho)$ is expressed as follows:

$$n(\rho) = \nu, \qquad \log (c_\nu/p_0) < \rho \leqq \log (c_{\nu+1}/p_0) \qquad (\nu = 1, 2, 3, \cdots).$$

From (27), $\log (c_\nu/p_0) = (1 + t_\nu)\nu \log \nu$, where $t_\nu \to 0$ as $\nu \to \infty$. Therefore

$$(1 + t_\nu)n(\rho) \log n(\rho) < \rho \leqq (1 + t_{\nu+1})[n(\rho) + 1] \log [n(\rho) + 1],$$

and

$$\frac{(1 + t_\nu) \log n(\rho)}{\log (1 + t_{\nu+1}) + \log [n(\rho) + 1] + \log \log [n(\rho) + 1]}$$

$$< \frac{\rho/\log \rho}{n(\rho)} < \frac{(1 + t_{\nu+1})[1 + 1/n(\rho)] \log [n(\rho) + 1]}{\log (1 + t_\nu) + \log n(\rho) + \log \log n(\rho)}.$$

The two end members of this inequality tend toward 1 as $\rho \to \infty$, and hence

$$n(\rho) \sim \frac{\rho}{\log \rho}. \tag{28}$$

From (28) it follows that the integral of Theorem I is comparable to $\int^\infty \frac{\log \rho \, d\rho}{\rho}$ which diverges. Consequently the condition for the parabolic case expressed in Theorem I is met by the Gamma surface. From (27) it follows that the series of Theorem II is comparable to the series $\sum_{\nu=1}^{\infty} \frac{\log \nu}{\nu}$ which is divergent, so that the condition of Theorem II is also met by the Gamma surface.

The same is true of Theorem III; for, from (25)

$$\lim_{\nu\to\infty} \frac{|A_\nu|}{\exp\left[\dfrac{\nu + 1}{\log (\nu + 1)}\right] \Big/ \log (\nu + 1)} = \lim_{\nu\to\infty} \frac{\Gamma(\nu - h_\nu)}{\exp\left[\dfrac{\nu + 1}{\log (\nu + 1)}\right]} = \infty.$$

Hence for ν sufficiently large

$$|A_\nu| > \frac{m}{\log (\nu + 1)} \exp\left[\frac{(\nu + 1)}{\log (\nu + 1)}\right],$$

where m is any positive constant, and so by Theorem III the surface is parabolic.

THE RICE INSTITUTE.

[11] See Nielsen, loc. cit., p. 91.

SOME PROPERTIES OF POLYNOMIAL SETS OF TYPE ZERO

By I. M. Sheffer

1. **Introduction.** Pincherle,[1] in his study of the difference equation

$$\sum_{n=1}^{k} c_n \phi(x + h_n) = f(x),$$

was led to consider a set of Appell polynomials, in infinite series of which solutions could be represented. We considered the same equation[2] by means of a different Appell set, the change resulting in a significant alteration of the regions of convergence (for the series). This permitted an enlargement of the class of functions $f(x)$ for which a solution could be shown to exist. Recently we treated the more general equation[3] (linear differential equation of infinite order)

$$L[y] \equiv a_0 y + a_1 y' + \cdots = f(x),$$

where, under suitable conditions on L and f, a solution was found. Here, too, it was possible to relate the equation to a corresponding problem of expanding functions in series of Appell polynomials. It is this close relation to functional equations that adds interest to the study of Appell sets.

As is well known, Appell sets $\{P_n(x)\}$ $(n = 0, 1, \cdots)$ are characterized by either of the equivalent conditions

$$P'_n(x) = P_{n-1}(x) \qquad (P_n \text{ a polynomial of degree } n); \tag{1.1}$$

$$A(t)e^{tx} \cong \sum_{0}^{\infty} P_n(x)t^n, \tag{1.2}$$

where $A(t) \cong \sum a_n t^n$ is a formal power series, and where the product on the left of (1.2) is formally expanded in a power series in accordance with the Cauchy rule. We shall say that the series $A(t)$ is the *determining series* for the set $\{P_n\}$.

For the particular equation

$$y(x + 1) - y(x) = f(x),$$

Pincherle used the Appell set with $A(t) = 1/(e^t - 1)$, getting essentially the Bernoulli polynomials. We used $A(t) = e^t - 1$, so that $n!P_n(x) = (x + 1)^n - x^n$. Now this equation is also associated with the important set of Newton polynomials

Received February 6, 1939; in preliminary form this paper was presented to the American Mathematical Society, April, 1936 under the title *A characterization of a class of polynomials*.

[1] Acta Mathematica, vol. 48(1926), pp. 279–304.

[2] Trans. Amer. Math. Soc., vol. 39(1936), pp. 345–379, and vol. 41(1937), pp. 153–159.

[3] This Journal, vol. 3(1937), pp. 593–609.

$$N_0(x) = 1, \qquad N_n(x) = \frac{x(x-1)\cdots(x-n+1)}{n!} \qquad (n = 1, 2, \cdots), \tag{1.3}$$

which is *not* an Appell set. Yet, it has properties analogous to those ((1.1) and (1.2)) of Appell polynomials. In fact,

$$\Delta N_n(x) \equiv N_n(x+1) - N_n(x) = N_{n-1}(x), \tag{1.4}$$

$$(1+t)^x = 1 \cdot e^{x \log(1+t)} = \sum_0^\infty N_n(x)t^n. \tag{1.5}$$

It is thus suggested that we define a class of *difference polynomial sets*, of which $\{N_n\}$ is a particular set, by means of the relations

$$\Delta P_n(x) = P_{n-1}(x) \qquad (n = 0, 1, 2, \cdots). \tag{1.6}$$

And more generally, we can use other operators than d/dx and Δ, to define further sets. We thus obtain all polynomial sets of *type zero* (as we denote them). The definition of sets of type zero generalizes readily to give sets of type one, two, $\cdots$, and of infinite type. (This is done immediately after relation (1.15).) This hierarchy of types is all-inclusive, in that every set of polynomials is of a definite type.

The main purpose of this paper is to bring to attention these sets of type zero and to indicate some of their properties. This section considers sets in general. §2 obtains various characterizations of zero type sets. Then, in §3, a study is made of the conditions on a set of zero type in order that it satisfy certain functional equations of finite order. As there are some known Tchebycheff sets that are of type zero, we next (§4) determine all zero type sets that are Tchebycheff sets. Lastly, in §5, we examine some extensions of the definition of type to type of higher order.

By a *set of polynomials* $\{P_n(x)\}$ $(n = 0, 1, 2, \cdots)$ we shall mean a sequence in which each P_n is of degree *exactly* n. We shall denote the set $\{P_n\}$ by P.

Lemma 1.1. *Let J be a linear operator applicable to the functions x^n $(n = 0, 1, \cdots)$ (and hence to all polynomials) and such that $J[x^n]$ is a polynomial of degree not exceeding n. Then J has the form*

$$J[y(x)] = \sum_0^\infty L_n(x)y^{(n)}(x), \tag{1.7}$$

valid for all polynomials, where $L_n(x)$ is a polynomial of degree not exceeding n.

To see this, define the $L_n(x)$ recurrently by the relations

$$J[x^n] = \sum_{k=0}^n L_k(x)\cdot n(n-1)\cdots(n-k+1)x^{n-k} \qquad (n = 0, 1, \cdots). \tag{1.8}$$

Since the degree of $J[x^n]$ does not exceed n, the degree of the L_n's are seen not to exceed their index. By construction, (1.7) now holds for $y(x) = x^n$, and therefore for all polynomials.

Of special interest is the case where $J[x^n]$ is always of degree $n - 1$.

LEMMA 1.2. *In order that the operator* (1.7) *carry every polynomial into one whose degree is less by precisely*[4] *one, it is necessary and sufficient that*

$$(1.9)\qquad L_0(x) = 0, \qquad L_n(x) = l_{n0} + l_{n1}x + \cdots + l_{n,n-1}x^{n-1} \qquad (n = 1, 2, \cdots)$$

and

$$(1.10)\qquad \lambda_n \equiv nl_{10} + n(n-1)l_{21} + \cdots + n!l_{n,n-1} \neq 0 \qquad (n = 1, 2, \cdots).$$

First suppose that $J[1] = 0$ and $J[x^n]$ is a polynomial of degree $n - 1$ $(n = 1, 2, \cdots)$. From (1.8) we find that the coefficients of x^n and x^{n-1} in $J[x^n]$ are respectively

$$l_{00} + nl_{11} + n(n-1)l_{22} + \cdots + n!l_{nn}, \qquad \lambda_n \qquad (n = 0, 1, \cdots).$$

Taking $n = 0, 1, \cdots$ successively, we see that $l_{ii} = 0$ $(i = 0, 1, \cdots)$, so that $L_n(x)$ is of degree less than n; and in order that $L[x^n]$ be of degree exactly $n - 1$, it is necessary that $\lambda_n \neq 0$. The conditions are thus necessary, and it is readily seen that they are also sufficient.

We shall assume without further mention that the operators with which we deal are of type (1.8) and that they fulfill the conditions of Lemma 1.2, so that they have the form

$$(1.11)\qquad J[y] = \sum_{n=1}^{\infty} (l_{n0} + \cdots + l_{n,n-1}x^{n-1})y^{(n)}(x)$$

with $\lambda_n \neq 0$ $(n = 1, 2, \cdots)$.

THEOREM 1.1. *Let P: $\{P_n(x)\}$ be a given set. There is a unique operator J for which*

$$(1.12)\qquad J[P_n] = P_{n-1} \qquad (n = 0, 1, \cdots).$$

If $y(x) = P_n$ $(n = 1, 2, \cdots)$ is substituted into (1.11), it is found that the l_{ij}'s exist to make (1.12) true and are uniquely determined. This is the assertion of the theorem. If P satisfies (1.12) we shall say that P corresponds to the operator J. Conversely, we have

THEOREM 1.2. *To each operator J correspond infinitely many sets P for which* (1.12) *holds. In particular, one and only one of these sets* (*which we call the basic set and denote by* $\{B_n\}$) *is such that*[5]

$$(1.13)\qquad B_0(x) = 1; \qquad B_n(0) = 0, \qquad n > 0.$$

If Q is any polynomial of degree s, it is found by direct substitution that a polynomial P exists, unique to within an additive constant, such that $J[P] = Q$;

[4] It is understood that $J[c] = 0$ for every constant c.

[5] The set $\{B_n\}$ is the "best approximation" set relative to the sequence of operators $J^0, J^1, J^2, \cdots$ according to the definition in the American Journal of Mathematics, vol. 57(1935), especially p. 593.

and P is of degree $s + 1$. Choosing $B_0(x) = 1$, we can then successively (and uniquely) determine $B_1, B_2, \cdots$ to satisfy $J[B_n] = B_{n-1}$, $B_n(0) = 0$ $(n > 0)$. Moreover, B_n is of degree exactly n. We thus have the existence of the *basic set*. That infinitely many sets exist is a consequence of the additive constant that is arbitrary. In fact, we have

COROLLARY 1.1. *A necessary and sufficient condition that P be a set corresponding to J is that there exist a sequence of numbers $\{a_n\}$ such that*

$$P_n(x) = a_0B_n(x) + a_1B_{n-1}(x) + \cdots + a_nB_0(x) \qquad (a_0 \neq 0). \tag{1.14}$$

Here the B_n's form the basic set for J.

If P satisfies (1.14), then P_n is of degree n, so that P is a set. Again,

$$J[P_n] = \sum a_iJ[B_{n-i}] = \sum a_iB_{n-i-1},$$

so that $J[P_n] = P_{n-1}$. This proves the sufficiency.

To establish the necessity, we first observe that constants $\{a_{ni}\}$ exist so that

$$P_n(x) = a_{n0}B_n(x) + \cdots + a_{nn}B_0(x).$$

From the relation $J[P_n] = P_{n-1}$ it follows that

$$a_{n0}B_{n-1} + \cdots + a_{n,n-1}B_0 = a_{n-1,0}B_{n-1} + \cdots + a_{n-1,n-1}B_0,$$

so that

$$a_{nj} = a_{n-1,j} \qquad (j = 0, 1, \cdots, n-1).$$

For a fixed j, this says that for every $n \geqq j$, all a_{nj}'s with second index j are equal. It is therefore permissible to drop the first index of each a_{nj}; which means that $\{a_j\}$ exists so that P_n is given by (1.14).

It can likewise be shown that

COROLLARY 1.2. *If P is a set corresponding to J, then a necessary and sufficient condition that $\{Q_n\}$ correspond to J is that constants $\{b_n\}$ exist so that*

$$Q_n = b_0P_n + b_1P_{n-1} + \cdots + b_nP_0. \tag{1.15}$$

DEFINITION. Let J be the (unique) operator corresponding to a given set P. *P is of type k* if in (1.11) no coefficient $L_n(x)$ is of degree exceeding k, but at least one is of degree k. If the degrees of the coefficients $L_n(x)$ are unbounded, then P is of *infinite type*.

From Theorem 1.2 follows

COROLLARY 1.3. *There are infinitely many sets for every type (finite or infinite).*

It is of interest to ask under what alterations, either of the set P itself or of the operator J which defines the type of P, the type is preserved. We consider two simple cases.

(i) Suppose P_n is replaced by c_nP_n, where $c_n \neq 0$ $(n = 0, 1, \cdots)$. Such a

transformation does not affect convergence properties of series in these polynomials, but it can very well change the type, as is readily established.

(ii) Let the operator

$$K[y] \equiv k_1 y' + k_2 y'' + \cdots \qquad (k_1 \neq 0) \tag{1.16}$$

be given. The following can be shown (analogously to Theorem 1.1): *If P is any set, there exists a unique operator J_K of form*

$$J_K[y] = \sum_{n=1}^{\infty} (l_{n0} + l_{n1}x + \cdots + l_{n,n-1}x^{n-1})K^n[y] \tag{1.17}$$

(where K^n means $K[K^{n-1}[y]]$), such that

$$J_K[P_n] = P_{n-1}. \tag{1.18}$$

Moreover,

$$\lambda_n \equiv nl_{10}k_1 + n(n-1)l_{21}k_1^2 + \cdots + n!l_{n,n-1}k_1^n \neq 0 \quad (n = 1, 2, \cdots). \tag{1.19}$$

There is also an analogue to Theorem 1.2.

Now suppose that we define the type of a set by the degrees of the polynomials $L_n(x)$ in (1.17). It is seen that no matter what operator K (of form (1.16)) is used, the type of P is the same.

2. **On sets of type zero.** Especially simple and important are sets of type zero. We shall find several characterizations for such sets. It will be convenient to restate the condition for a set of type zero as follows: *P is of type zero if*

$$J[P_n] = P_{n-1} \qquad (n = 0, 1, 2, \cdots), \tag{2.1}$$

where

$$J[y] = c_1 y' + c_2 y'' + c_3 y''' + \cdots \qquad (c_1 \neq 0). \tag{2.2}$$

DEFINITION. The formal series

$$J(t) \cong c_1 t + c_2 t^2 + c_3 t^3 + \cdots \tag{2.3}$$

will be called the *generating series* (*or function*) for the operator (2.2).

That Appell sets are of type zero follows from the fact that the generating series is $J(t) = t$. Similarly, for Newton polynomials (and for all the *difference sets*—cf. (1.6)), $J(t) = e^t - 1$.

Let P be of type zero, corresponding to the operator J. Let the formal power series inverse of (2.3) be

$$H(t) \cong s_1 t + s_2 t^2 + \cdots \qquad (s_1 = c_1^{-1} \neq 0), \tag{2.4}$$

obtained from[6]

$$J(H(t)) \cong H(J(t)) \cong t. \tag{2.5}$$

If e is raised to the power $xH(t)$, the expression will have a formal power series expansion in t, in which the coefficient of t^n involves only $s_1, \cdots, s_n$. On multiplying by the formal power series[7]

$$A(t) \cong \sum_0^\infty a_n t^n \qquad (a_0 \neq 0), \tag{2.6}$$

a new series (in t) is obtained, in which the coefficient of t^n now involves only $a_0, \cdots, a_n; s_1, \cdots, s_n$. In fact, this coefficient is a polynomial in x of degree n, and we have, furthermore,

THEOREM 2.1. *A necessary and sufficient condition that P be of type zero corresponding to the operator J of (2.2) is that $\{a_n\}$ exist so that*

$$A(t)e^{xH(t)} \cong \sum_{n=0}^\infty P_n(x)t^n. \tag{2.7}$$

From (1.14) of Corollary 1.1 it is seen that both the necessary and sufficient parts will follow if we can show that for the basic set $\{B_n\}$ (corresponding to J) we have

$$e^{xH(t)} \cong \sum B_n(x)t^n. \tag{2.8}$$

Let $\exp\{xH(t)\}$ have the expansion $\sum C_n(x)t^n$. Then $C_n(x)$ is a polynomial of degree exactly n. On setting $x = 0$, we obtain $1 \cong \sum C_n(0)t^n$, so that $C_0(0) = C_0(x) = 1$, $C_n(0) = 0$ $(n > 0)$. By Theorem 1.2, $\{C_n\}$ will therefore be the basic set if we establish the relation $J[C_n] = C_{n-1}$ $(n = 0, 1, \cdots)$. Operate on the $C_n(x)$-series with J. This gives

$$\sum_0^\infty J[C_n]t^n \cong J[\exp\{xH(t)\}] \cong \{c_1H + c_2H^2 + \cdots\}\cdot\exp\{xH\}$$

$$\cong J(H(t))\cdot\exp\{xH\} \cong t\cdot\exp\{xH\} \cong \sum_0^\infty C_{n-1}(x)t^n;$$

[6] If the series for $J(t)$ is formally substituted for t in (2.4), and coefficients combined (in the usual way) to form a single power series in t, the coefficient of t^n is for each n a polynomial in $c_1, c_2, \cdots, c_n, s_1, \cdots, s_n$. It is possible to choose s_n recurrently and uniquely as a simple function of $c_1, \cdots, c_n, s_1, \cdots, s_{n-1}$, so that the power series reduces to the single term t. This sequence of s_i's is the one to be used in (2.4).

[7] The condition $a_0 \neq 0$ is to insure that $P_n(x)$ in (2.7) is of degree n and not less. But $a_0 \neq 0$ is no essential restriction. See, for example, the footnote on page 916 of Bull. Amer. Math. Soc., vol. 41(1935).

and on comparing like powers of t, we obtain $J[C_n] = C_{n-1}$. Thus, (2.8), holds.[8]

Corollary 2.1. *In* (2.7) *the numbers* $\{a_n\}$ *are the same as in* (1.14).

Thus, for Appell sets, (2.7) holds with $H(t) = t$ (cf. (1.2)), and for difference sets (including the Newton polynomials), $H(t) = \log(1 + t)$, so that *a necessary and sufficient condition for a difference set P is that*

$$A(t)(1 + t)^x \cong \sum P_n(x)t^n. \tag{2.9}$$

Another familiar set of polynomials of type zero is given by the Laguerre polynomials which satisfy the relation

$$\frac{1}{1 - t} \cdot \exp\left\{\frac{-tx}{1 - t}\right\} = \sum L_n(x)t^n. \tag{2.10}$$

For this set,

$$A(t) = \frac{1}{1 - t}, \qquad H(t) = J(t) = \frac{-t}{1 - t} = -\sum_0^\infty t^{n+1},$$

so that

$$L_{n-1}(x) = -(L_n' + L_n'' + L_n''' + \cdots).$$

[8] Some words are in order regarding the validity of the above proof, which uses *formal series*; particularly, since like arguments (as well as obvious modifications) will be used again. Let k be any positive integer. Consider the operator

$$J_k[y] = c_1y' + \cdots + c_ky^{(k)}$$

and the generating series

$$J_k(t) = c_1t + \cdots + c_kt^k.$$

Let $H_k(t)$ be the inverse of $J_k(t)$:

$$J_k(H_k(t)) = H_k(J_k(t)) = t,$$

and define $C_{kn}(x)$ by the *convergent* expansion

$$\exp\{xH_k(t)\} = \sum_{n=0}^\infty C_{kn}(x)t^n.$$

The argument advanced above is now completely legitimate, giving the relations

$$J_k[C_{kn}(x)] = C_{k,n-1}(x).$$

Now it is readily seen that the series for $H(t)$ and for $H_k(t)$ agree through the term in t^k, whence the same is true for the two series for $\exp\{xH(t)\}$ and $\exp\{xH_k(t)\}$. This means that $C_{kn}(x) = C_n(x)$ $(n = 0, 1, \cdots, k)$. As k is arbitrary, it follows that $J[C_n] = C_{n-1}$ for *all n*. This establishes (2.8) and, therefore, (2.7).

Having thus shown that the use of formal power series yields the correct result (in the above case), we shall not hesitate to use such power series in what follows, leaving the argument in the present footnote as a guide to further "validity proofs".

Every set satisfies infinitely many linear functional equations. One of the simplest for sets of zero type is given by

THEOREM 2.2. *Let P be of type zero corresponding to operator J, and let $A(t)$ be its determining series. Then P satisfies the equation*

$$L[y(x)] \equiv \sum_{k=1}^{\infty} (q_{k0} + xq_{k1})J^k[y] = \lambda y, \tag{2.11}$$

where $\lambda = n$ for $y = P_n(x)$. The q's are defined by

$$\frac{A'(t)}{A(t)} \cong \sum_{0}^{\infty} q_{n+1,0}t^n, \tag{2.12}$$

$$H'(t) \cong \sum_{0}^{\infty} q_{n+1,1}t^n. \tag{2.13}$$

Suppose each side of (2.11) (with $y = P_n$, $\lambda = n$) is multiplied by t^n and a summation made from $n = 0$ to $n = \infty$. There result two power series in t. (2.11) will be established if we show that these series are formally equal. Now the right-hand series is

$$t \sum nP_n t^{n-1} \cong t\frac{d}{dt}\{Ae^{xH}\} \cong te^{xH}\{A' + xH'A\}.$$

Also,

$$\sum J^k[P_n]t^n \cong t^k \sum P_n t^n \cong t^k A e^{xH},$$

so that the left-hand series is

$$\sum_{n=0}^{\infty} \sum_{k=1}^{\infty} (q_{k0} + xq_{k1})J^k[P_n]t^n \cong \sum_{k=1}^{\infty} (q_{k0} + xq_{k1})t^k A e^{xH},$$

and if we use (2.12) and (2.13), this becomes

$$Ae^{xH}\left\{t\frac{A'}{A} + xtH'\right\}.$$

Hence, the two series are equal, and (2.11) holds.

Since (2.11) is linear and homogeneous, multiplication of a solution by a constant again yields a solution. But such multiplication may destroy the property of being a zero type set. We cannot therefore obtain a complete converse to Theorem 2.2. But we do have

COROLLARY 2.2. *Given an operator J. If a set P satisfies an equation of form* (2.11), *where $\lambda = n$ for $y = P_n(x)$, and if $\{q_{n1}\}$ is related to J by* (2.13), *then non-zero constants $\{h_n\}$ exist so that $\{h_n P_n\}$ is a set of type zero corresponding to J. Its determining series $A(t)$ is then given by* (2.12).

For, define $A(t)$ by (2.12), the arbitrary multiplicative constant which enters being given any non-zero value. By Theorem 2.2, the set $\{R_n\}$, of type zero, corresponding to J and with determining series $A(t)$, satisfies (2.11). Now, it is

readily found that for $\lambda = n$ equation (2.11) has a polynomial solution, and that this polynomial is unique to within an arbitrary multiplicative constant. Hence, $\{h_n\}$ exists so that $R_n = h_n P_n$.

As a characterization of sets of type zero, Theorem 2.2 is not wholly satisfactory, since it involves the operator J of the set. This objection is removed in

THEOREM 2.3. *If P is of type zero, it satisfies an equation of the form*

$$M[y(x)] \equiv \sum_{k=1}^{\infty} (r_{k0} + xr_{k1})y^{(k)}(x) = \lambda y(x), \tag{2.14}$$

where $\lambda = n$ for $y = P_n$. Moreover, the operator J and determining function A corresponding to P are related to the r's as follows:

$$\sum_{k=1}^{\infty} r_{k0}t^k \cong \frac{uA'(u)}{A(u)}, \tag{2.15}$$

$$\sum_{k=1}^{\infty} r_{k1}t^k \cong uH'(u), \tag{2.16}$$

where $u = J(t)$. Conversely, if a set P satisfies equation (2.14), *then non-zero constants $\{h_n\}$ exist such that $\{h_nP_n\}$ is of type zero.*

To see this we observe that if in (2.11) we write out each $J^k[y]$ as a series of derivatives of y and collect all terms with the same order of derivative, then to each k there are only a finite number of terms in $y^{(k)}(x)$. The result of this collecting of terms is to give us the equivalent equation (2.14).

If in (2.11) and (2.14) we replace each derivative $y^{(k)}(x)$ by t^k, we obtain of course the same formal series (since (2.14) is merely a regrouping of terms in (2.11)). That is,

$$\sum_{k=1}^{\infty} (r_{k0} + xr_{k1})t^k \cong \sum_{k=1}^{\infty} (q_{k0} + xq_{k1})J^k(t).$$

On using (2.12) and (2.13), we obtain (2.15) and (2.16). The converse follows as in Corollary 2.2.

One obtains a generalization by replacing $y^{(k)}(x)$ in $M[y]$ by $K^k[y]$, where K is an operator of form (1.16). This yields the following theorem (proved as was Theorem 2.3):

THEOREM 2.4. *Let operator K be given. If P is a set of type zero, it satisfies an equation of form*[9]

$$T[y(x)] \equiv \sum_{k=1}^{\infty} (r_{k0} + xr_{k1})K^k[y] = \lambda y(x), \tag{2.17}$$

where $\lambda = n$ for $y = P_n$. The operator J and determining function A corresponding to P are related to the r's as follows:

[9] We observe, on comparing the coefficient of x^n on both sides of (2.17), that $r_{11} = k_1^{-1}$.

$$\sum_{k=1}^{\infty} r_{k0}[K(t)]^k \cong \frac{uA'(u)}{A(u)}, \tag{2.18}$$

$$\sum_{k=1}^{\infty} r_{k1}[K(t)]^k \cong uH'(u), \tag{2.19}$$

$$u = J(t).$$

Conversely, if a set P satisfies equation (2.17), *then non-zero constants* $\{h_n\}$ *exist so that* $\{h_n P_n\}$ *is of type zero.*

From (2.11) of Theorem 2.2 there follows a further characterization of sets of type zero, expressed solely in terms of the members of the set itself. It is

THEOREM 2.5. *A necessary and sufficient condition that a set P be of type zero is that constants q_{k0}, q_{k1} exist so that*

$$\sum_{k=1}^{\infty} (q_{k0} + xq_{k1})P_{n-k}(x) = nP_n(x) \qquad (n = 1, 2, \cdots). \tag{2.20}$$

The operator J and the determining series A for P are related to the q's by (2.12) *and* (2.13).

Let (2.7) be differentiated with respect to x. On equating coefficients of like powers of t, we obtain

$$P_n'(x) = s_1 P_{n-1}(x) + s_2 P_{n-2}(x) + \cdots + s_n P_0(x) \qquad (n = 1, 2, \cdots), \tag{2.21}$$

whence we have

THEOREM 2.6. *A necessary and sufficient condition that a set P be of type zero is that constants* $\{s_n\}$ *exist for which* (2.21) *holds; in this case the operator J corresponding to P is determined through* $\{s_n\}$ *by means of* (2.4) *and* (2.5).[10]

Theorem 2.6 will later be seen to generalize to sets of all types. (Cf. Lemma 5.1.) It is of interest to compare (2.20) and (2.21). The latter involves only J (through H), so that *all* zero type sets for one and the same operator satisfy the *same* equation of form (2.21). On the other hand, both J and A are involved by the constants present in (2.20), so that if sets $\{P_n\}$ and $\{Q_n\}$ both satisfy (2.20), then there is a $c \neq 0$ such that $Q_n = cP_n$, $n > 1$; i.e., there is an essentially unique set satisfying (2.20) for given q_{k0}, q_{k1}.

If (2.7) is differentiated with respect to x, the left side is multiplied by $H(t)$, and P_n' replaces P_n on the right. Recalling that H begins with a term in t, we see that $Q_n = P_{n+1}'$ is a set of zero type, corresponding to the same operator J as does P. In other words, we have

THEOREM 2.7. *If P is of type zero, then so are the sets* $\{P_{n+1}'\}$, $\{P_{n+2}''\}$, $\{P_{n+3}'''\}$, $\cdots$ *; and they all correspond to the same operator as does P. More*

[10] A simple extension of Theorem 2.6 is the following: *Let P be a set with operator J whose inverse is H, and let K be an operator of form* (1.16). *Set*

$$K(H(t)) \cong \alpha_1 t + \alpha_2 t^2 + \cdots.$$

Then

$$K[P_n(x)] = \alpha_1 P_{n-1}(x) + \alpha_2 P_{n-2}(x) + \cdots + \alpha_n P_0(x) \qquad (n = 1, 2, \cdots)$$

generally, let K be of form (1.16). *Then, if P is of type zero, so are* $\{K[P_{n+1}]\}$, $\{K^2[P_{n+2}]\}$, $\cdots$, *and they correspond to the same operator as does P.*

Let us apply the preceding characterizations to some well-known zero-type sets.

Example 1. $P_n(x) = x^n/n!$. Then

$$\sum_0^\infty P_n(x)t^n = e^{tx}, \qquad A(t) = 1, \qquad H(t) = J(t) = t.$$

It is readily determined that (2.11) and (2.14) become

$$xP_n' = nP_n,$$

and (2.20), (2.21) become, respectively,

$$xP_{n-1} = nP_n, \qquad P_n' = P_{n-1}.$$

Also, (2.17) holds with $r_{k0} = 0$ and r_{k1} determined from

$$\sum_{k=1}^\infty r_{k1}[K(t)]^k = t.$$

Example 2. Laguerre polynomials $\{L_n(x)\}$. Here

$$\sum_0^\infty L_n(x)t^n = \left(\frac{1}{1-t}\right)\exp\left\{\frac{-xt}{1-t}\right\}, \qquad A(t) = \frac{1}{1-t}, \qquad J(t) = H(t) = \frac{-t}{1-t}.$$

It is found that (2.11) and (2.14) become

$$\sum_{k=1}^\infty (1 - kx)J^k[L_n(x)] = nL_n(x) \qquad (J[y] \equiv -y' - y'' - y''' - \cdots),$$

$$(x-1)L_n' - xL_n'' = nL_n,$$

while (2.20), (2.21) reduce to

$$\sum_{k=1}^\infty (1-kx)L_{n-k}(x) = nL_n(x), \qquad L_n'(x) = -[L_{n-1}(x) + L_{n-2}(x) + \cdots].$$

Most of these relations are known. In (2.17), r_{k0}, r_{k1} are determined by the series

$$\sum_{k=1}^\infty r_{k0}[K(t)]^k = -t, \qquad \sum_{k=1}^\infty r_{k1}[K(t)]^k = t - t^2.$$

Example 3. Hermite polynomials $\{H_n(x)\}$. Here[11]

$$\sum_{n=0}^\infty H_n(x)t^n = \exp\{-t^2 + 2tx\}, \qquad A(t) = e^{-t^2}, \qquad H(t) = 2t, \qquad J(t) = \tfrac{1}{2}t.$$

[11] It is more common to define $H_n(x)$ so that $H_n/n!$ is the coefficient of t^n. We find it simpler to adopt the present definition.

Relations (2.11) and (2.14) become

$$2xH_n' - H_n'' = 2nH_n ,$$

and (2.20), (2.21) become

$$2xH_{n-1} - 2H_{n-2} = nH_n , \qquad H_n' = 2H_{n-1} .$$

These are also well known. The defining relations for r_{k0}, r_{k1} in (2.17) are

$$\sum_{k=1}^{\infty} r_{k0}[K(t)]^k = -\tfrac{1}{4}t^2, \qquad \sum_{k=1}^{\infty} r_{k1}[K(t)]^k = t.$$

Example 4. Newton Polynomials. Here

$$\sum_0^{\infty} N_n(x)t^n = (1+t)^x, \qquad A(t) = 1, \qquad H(t) = \log(1+t), \qquad J(t) = e^t - 1.$$

It is seen that (2.11), (2.14), (2.20), (2.21) become, respectively,

$$x \sum_{k=1}^{\infty} (-1)^{k-1}\Delta^k N_n(x) = nN_n(x), \qquad x \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k!} N_n^{(k)}(x) = nN_n(x),$$

$$x \sum_{k=1}^{\infty} (-1)^{k-1} N_{n-k}(x) = nN_n(x),$$

$$N_n'(x) = N_{n-1}(x) - \frac{1}{2}N_{n-2}(x) + \cdots + \frac{(-1)^{n-1}}{n} N_0(x).$$

In (2.17), $r_{k0} = 0$ and r_{k1} is determined by

$$\sum_{k=1}^{\infty} r_{k1}[K(t)]^k = 1 - e^{-t}.$$

3. **On sets of zero type satisfying finite order equations.** In the Bulletin paper (cited in footnote 7) those Appell sets were determined that satisfy a finite order linear differential equation with polynomial coefficients. Here we extend the problem to the case of a finite order equation in an operator K of form (1.16), satisfied by a zero type set P corresponding to the operator J and determining function A. We first restrict our attention to equations of form (2.17).

In order that a set P, corresponding to a given J and A, satisfy a finite order equation of form (2.17):

$$T[y(x)] \equiv \sum_{k=1}^{s} (r_{k0} + xr_{k1})K^k[y] = \lambda y(x), \tag{3.1}$$

with $\lambda = n$ for $y = P_n(x)$, it is necessary and sufficient that the following relations hold:[12]

[12] It was shown in a footnote to Theorem 2.4 that $r_{11} = k_1^{-1}$.

$$\sum_{k=1}^{m} r_{k0}[K(t)]^k = F_0(t), \tag{3.2}$$

$$\sum_{k=1}^{q} r_{k1}[K(t)]^k = F_1(t), \tag{3.3}$$

where $s = \max(m, q)$ and F_0, F_1 are defined by

$$F_0(t) = \{uA'(u)/A(u)\}, \qquad F_1(t) = \{uH'(u)\} \qquad (u = J(t)). \tag{3.4}$$

Suppose P satisfies (3.1). Since $r_{11} = 1/k_1 \neq 0$, at least one r_{k1} is not zero, and we let r_{q1} be the last non-zero one. On the other hand, all the r_{k0}'s may be zero. This is true, as we see from (3.2), if and only if $A(t)$ is a constant. (3.3) then gives us

THEOREM 3.1. *If P corresponds to a given J and A, with $A(t) \equiv c \neq 0$, then a necessary and sufficient condition that P satisfy a finite order equation of form* (3.1) *is that $F_1(t)$ be a polynomial in $K(t)$.*

Now suppose that $A \not\equiv c$. Then the r_{k0}'s are not all zero, and we let r_{m0} be the last non-zero one. Since (3.2) and (3.3) are polynomials having a common root $K(t)$, their Sylvester determinant vanishes:

$$\begin{array}{r} q \text{ rows}\left\{\begin{array}{l} \\ \\ \\ \end{array}\right. \\ m \text{ rows}\left\{\begin{array}{l} \\ \\ \\ \end{array}\right. \end{array}
\begin{vmatrix} r_{m0}\ r_{m-1,0} & \cdots & r_{10} - F_0(t) & \cdots & \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ & \cdots & r_{m0}\ r_{m-1,0} & \cdots & r_{10} - F_0(t) \\ r_{q1}\ r_{q-1,1} & \cdots & r_{11} - F_1(t) & \cdots & \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ & \cdots & r_{q1}\ r_{q-1,1} & \cdots & r_{11} - F_1(t) \end{vmatrix} = 0. \tag{3.5}$$

(In (3.5), elements that are zero are not indicated.)

Conversely, suppose (3.5) holds for a choice of the r's, with $r_{m0} \neq 0$, $r_{q1} \neq 0$, and let $s = \max(m, q)$. Then there is a common solution $K(t)$ of (3.2) and (3.3), which solution can be expressed as a formal power series in t. Now from (1.16), K must begin with a term in t. Suppose the common solution $K(t)$ does not have this property. Then let us set $K(t) = c + K_1(t)$, where $K_1(0) = 0$. On substituting into (3.2) and (3.3), we obtain two polynomials (on the left) with constant terms c_0, c_1, say. Now for $t = 0$, since $F_0(0) = F_1(0) = 0$, we get $c_0 = c_1 = 0$. Hence, $K_1(t)$ is also a common solution, and it is zero at $t = 0$. We shall then use K_1, or, what amounts to the same thing, we may suppose that $K(0) = 0$ to begin with. If K begins with a term of higher degree than one in t, then so does the left side of (3.3), and therefore the right side. But $tH'(t)$ begins with a term in t, as does J; hence, so does $F_1(t)$. As this is a contradiction, it follows that $K(t)$ begins with a term in t, and thus the common solution K has the properties of (1.16). Therefore P satisfies (3.1). This establishes

THEOREM 3.2. *Let P correspond to J and A, where $A \not\equiv c$. A necessary and sufficient condition that P satisfy a finite order equation of form* (3.1) *is that $F_0(t)$ and $F_1(t)$ satisfy* (3.5) *for some choice of r_{k0}, r_{k1} (with $r_{m0} \neq 0$, $r_{q1} \neq 0$).*

As (3.5) is to be an identity in t, it will likewise be an identity in u where $u = J(t)$ (so that $t = H(u)$). Accordingly we have

COROLLARY 3.1. *In Theorem* 3.2 *the functions* $F_0(t)$, $F_1(t)$ *may be replaced by* $uA'(u)/A(u)$, $uH'(u)$.

Theorem 3.1 can be generalized to

THEOREM 3.3. *If in Theorem* 3.1 *the condition* "$A(t) \equiv c \neq 0$" *is replaced by the condition*

$$F_0(t) = \sum_{k=1}^{p} \sigma_p[F_1(t)]^p \qquad (\sigma_p \neq 0), \tag{3.6}$$

then the same conclusion holds.

For, $K(t)$ can be defined by (3.3), whereupon $F_0(t)$ is expressible as a polynomial in $K(t)$. That is, an equation of form (3.2) holds. Hence, P satisfies (3.1).

There are infinitely many pairs of functions F_0, F_1 satisfying a given relation (3.5). One can, for example, give *one* of F_0, F_1 arbitrarily. This suggests examination of the following question: Given one of the three elements J, A, K, to what extent are the others determined so that (3.1) holds?

Case I. *Given* $J(t)$. This determines $H(t)$ and therefore $F_1(t)$. In fact, one easily finds from (2.5) that

$$F_1(t) = J(t) \div J'(t). \tag{3.7}$$

DEFINITION. Given a function (or formal series) $f(t)$, beginning with a term in t. We denote by $\mathfrak{F}\{f(t)\}$ the class of all (formal) series $z(t)$, beginning with a term in t, satisfying a relation of the form

$$p_1z + p_2z^2 + \cdots + p_nz^n = f(t), \tag{3.8}$$

the p's being constants, with $p_1 \neq 0$. $\mathfrak{F}\{f(t)\}$ thus represents a special class of algebraic functions of $f(t)$.

In terms of this definition, we see that K is determined from (3.3) as a member of the class $\mathfrak{F}\{F_1(t)\}$. That is, J being given, K must be in $\mathfrak{F}\{F_1(t)\}$, but can be an arbitrary member of this class. Consider any such $K(t)$. From $u = J(t)$ follows $t = H(u)$, so that $K(t) = K(H(u)) = K^*(u)$. $A(u)$ is then determined by the (necessary and sufficient) condition that $uA'(u)/A(u)$ be a polynomial in $K^*(u)$. Observing the wide choice possible for K, after which a further wide choice for A exists, we see that to each operator J correspond a large variety of polynomial sets P satisfying an equation of form (3.1).

Case II. *Given* $K(t)$. Then $F_1(t)$ is to be a polynomial in $K(t)$, lacking a constant term and with coefficient of the linear term equal to k_1^{-1}. For all such F_1, we determine $J(t)$ from (3.7). Having now K and J, we obtain A as in Case I.

Case III. *Given* $A(t)$. Let $K^*(u)$ be any member of the class $\mathfrak{F}\{uA'(u)/A(u)\}$. Determine $uH'(u)$ as any polynomial in $K^*(u)$ beginning with a linear

term, and solve for $H(u)$. Relation (2.5) gives us $J(u)$, and with this J we determine $K(t)$ from the identity $K(t) = K^*(J(t))$.

As illustration, choose $K(t) \equiv t$, so that (3.1) is a differential equation. From Case II we see that

$$\frac{J'}{J} = \frac{1}{tQ(t)},$$

where $Q(t)$ is any polynomial of form

$$Q(t) = 1 + l_1 t + \cdots + l_q t^q, \tag{3.9}$$

and that therefore

$$J(t) = ct \cdot \exp\left\{\int_0^t \frac{1 - Q(t)}{tQ(t)}\, dt\right\}. \tag{3.10}$$

The inverse function H can be found from the power series for (3.10) or from the differential equation[13]

$$uH'(u) = H(u)\{1 + l_1 H(u) + \cdots + l_q H^q(u)\}, \tag{3.11}$$

with the condition that $H(u)$ begins with the term t/c. And finally, A is obtained as the solution of the differential equation

$$\frac{uA'(u)}{A(u)} = b_1 H(u) + \cdots + b_m H^m(u),$$

where the b's are arbitrary, but $b_1 \neq 0$; i.e.,

$$A(u) = \gamma \cdot \exp\left\{\int_0^u \frac{1}{u}\,[b_1 H(u) + \cdots + b_m H^m(u)]\, du\right\}, \tag{3.12}$$

γ = arbitrary constant.

Relations (3.10), (3.11) *and* (3.12) *are thus necessary and sufficient conditions that equation* (3.1) *be satisfied for* $K(t) = t$.

It has already been remarked that (2.17) is not the only linear functional equation satisfied by a set P. In fact, extending a result in the Bulletin paper (loc. cit., p. 914), it is easy to show that given an operator K of form (1.16), and given any set P (which need not be of type zero), polynomials $\{L_n(x)\}$ with L_n of degree $\leqq n$, and characteristic numbers $\{\lambda_n\}$, can be chosen, and indeed in infinitely many ways, so that the set P is a solution of the equation[14]

$$L[y(x)] \equiv \sum_{n=1}^{\infty} L_n(x) K^n[y] = \lambda y \tag{3.13}$$

(with $\lambda = \lambda_n$ for $y = P_n$).

[13] $K(t) = t = K^*(u) = K(H(u))$. Therefore, $H(u) = t$, and $K^*(u) = H(u)$. (3.11) then follows from $tQ(t) = uH'(u)$ $(u = J(t))$ if we write $t = H(u)$.

[14] It is no restriction to have the summation begin with $n = 1$, for if a term $n = 0$ is present, it is of the form $L_0 y = cy$, and this can be absorbed into the right side. The only effect is to alter all the λ_n's by the amount $-c$.

An equation (3.13) is said to be of *finite order* r if $L_n(x) \equiv 0$ for $n > r$, but $L_r \not\equiv 0$. We now investigate conditions under which a set P of type zero satisfies a finite order equation of form (3.13).

LEMMA 3.1. *A set P (not necessarily of type zero) cannot satisfy two different equations of form* (3.13) *if the characteristic numbers are respectively the same.*

For, suppose P satisfies (3.13) and also $L^*[y] = \lambda y$ (whose coefficients and characteristic numbers are $L_n^*(x)$ and $\lambda_n^* = \lambda_n$). By subtraction,

$$\sum_{n=1}^{\infty} (L_n - L_n^*)K^n[P_s] = 0 \qquad (s = 1, 2, \cdots).$$

Now, $K^n[P_s] = 0$, $n > s$, and $K^n[P_n] = \text{constant} \neq 0$. Hence, on setting $s = 1, 2, \cdots$ successively, we find that $L_s - L_s^* = 0$, $s \geqq 1$. The two equations, supposedly different, are thus identical.

If we set

$$L_n(x) = l_{n0} + \cdots + l_{nn}x^n, \tag{3.14}$$

the characteristic numbers λ_n of (3.13) are given by

$$\lambda_n = nk_1l_{11} + n(n-1)k_1^2l_{22} + \cdots + n!k_1^nl_{nn} \ (n = 0, 1, \cdots). \tag{3.15}$$

This is seen on equating the coefficient of x^n on both sides of $L[P_n] = \lambda_n P_n$.

Suppose P is a set of zero type. We know that it satisfies (2.17), which is a particular case of (3.13):

$$T[P_n] \equiv \sum_{k=1}^{\infty} S_{1k}(x)K^k[P_n] = nP_n, \tag{3.16}$$

where

$$S_{1k}(x) = r_{k0} + xr_{k1}, \qquad r_{11} = 1/k_1 \neq 0. \tag{3.17}$$

Now define operators T_k by

$$T_k[y] \equiv T(T-1)(T-2)\cdots(T-k+1)[y]. \tag{3.18}$$

THEOREM 3.4. *If the zero type set P satisfies* (3.13), *then* (3.13) *can be expressed in the canonical form*

$$L[y] \equiv \sum_{n=1}^{\infty} \alpha_n T_n[y] = \lambda y, \tag{3.19}$$

where T_n is given by (3.18) *and α_n by*

$$\alpha_n = l_{nn}\cdot k_1^n . \tag{3.20}$$

To see this, denote the operator of (3.19) by $L^*[y]$. If in (2.17) $K[y]$ and its iterates are replaced by series in derivatives of y, using (1.16), we can write $T[y]$ as[15]

$$\text{(a)} \qquad T[y] = \sum_{k=1}^{\infty} Q_{1k}(x) y^{(k)}(x),$$

where Q_{1k} is a polynomial of degree $\leqq 1$. Iteration gives

$$\text{(b)} \qquad T^n[y] = \sum_{k=1}^{\infty} Q_{nk}(x) y^{(k)}(x),$$

where Q_{nk} is of degree $\leqq \max(n, k)$. From this it follows that

$$\text{(c)} \qquad T_n[y] = \sum_{k=1}^{\infty} R_{nk}(x) y^{(k)}(x),$$

R_{nk} being of degree $\leqq \max(n, k)$.

On replacing each $y^{(k)}(x)$ by its equivalent as a series of iterates of $K[y]$, (c) becomes[16]

$$\text{(d)} \qquad T_n[y] = \sum_{k=1}^{\infty} S_{nk}(x) K^k[y],$$

where S_{nk} is a polynomial of degree $\leqq \max(n, k)$. Since from (3.16) $T[P_n] = nP_n$, therefore

$$\text{(3.21)} \qquad T_n[P_k] = k(k-1) \cdots (k-n+1) P_k,$$

and in particular,

$$\text{(3.22)} \qquad T_n[P_k] = 0, \qquad n > k.$$

Using (3.22) in (d) for $k = 1, 2, \cdots, n-1$, we find that $S_{nk} = 0$, $k < n$, so that (d) can be written

$$\text{(3.23)} \qquad T_n[y] = \sum_{k=n}^{\infty} S_{nk}(x) K^k[y].$$

If we substitute this expression into $L^*[y]$ (given by (3.19)) and collect like iterates of K, we obtain for L^* the form

$$L^*[y] = \sum_{k=1}^{\infty} \{\alpha_1 S_{1k} + \cdots + \alpha_k S_{kk}\} K^k[y].$$

That is, L^* can be written in the form of (3.13). By Lemma 3.1 it will follow that L^* and L are identical if we show that the respective characteristic numbers are the same. For L the numbers are λ_n, given by (3.15). From (3.21) we

[15] Since $K(t)$ begins with a term in t, and therefore $K^s(t)$ with a term in t^s, the coefficient of any $y^{(k)}(x)$ in (a) is obtained from only a *finite number* of coefficients of (2.17). Hence, the coefficients in (a) are well-determined.

[16] The point of the preceding footnote applies here also.

see that λ_n^* (for L^*) is given by

$$\lambda_n^* = \sum_{k=1}^{\infty} \alpha_k \cdot n(n-1) \cdots (n-k+1),$$

and this is precisely λ_n. The theorem is thus established.

COROLLARY 3.2. *Under the conditions of Theorem* 3.4, *the coefficients* $L_n(x)$ *of* (3.13) *are given by*

$$L_n(x) = \alpha_1 S_{1n}(x) + \cdots + \alpha_n S_{nn}(x) \qquad (n = 1, 2, \cdots). \tag{3.24}$$

To justify the phrase "canonical form", it should be shown that every equation of form (3.19) has a solution of zero type. That is,

THEOREM 3.5. *Let K be an operator of form* (1.16), *and let* $T[y]$ *be of form*

$$T[y] \equiv \sum_{k=1}^{\infty} (r_{k0} + xr_{k1})K^k[y]$$

with $r_{11} = 1/k_1 \neq 0$. *Then for every choice of* α*'s* (*not all zero*), *the equation*

$$L[y] \equiv \sum_{n=1}^{\infty} \alpha_n T_n[y] = \lambda y$$

is satisfied by a zero type set.

In fact, $T[y]$ serves to define a zero type set P by virtue of Theorem 2.4 and the relation (2.17). This same set P will clearly satisfy the above equation $L[y] = \lambda y$, and this is what was to be shown.[17]

LEMMA 3.2. *In order that a zero type set P satisfy a finite order equation of form* (3.13) *it is necessary and sufficient that in the canonical form* (3.19) (*into which* (3.13) *can be cast*) *the following two conditions hold*:

$$\begin{cases} \alpha_n = 0, & n > r; \\ \alpha_1 S_{1n}(x) + \cdots + \alpha_r S_{rn}(x) = 0, & n > r. \end{cases} \tag{3.25}$$

If r is the smallest positive integer for which this is true, the equation is of order r.

For, if (3.13) is of order r, then from (3.20), $\alpha_n = 0$, $n > r$; and from (3.24) the other relation of (3.25) follows. Conversely, suppose (3.25) holds. Then (3.24) yields the relations $L_n(x) = 0$, $n > r$. The assertion as to the order is obvious.

The function $K(t)$ can be expanded in a power series in $J(t)$, where J is the operator for set P:

[17] However, it cannot be asserted here (as was the case in Theorem 2.4) that if Q is any set satisfying $L[Q_n] = \lambda_n Q_n$, then there exist non-zero constants c_n such that $P_n = c_n Q_n$ is of type zero. For, it may now happen that two or more λ's are equal. Suppose $\lambda_m = \lambda_n$, which value we call λ'. Q_m and Q_n are solutions of $L[y] = \lambda' y$, and therefore so is $aQ_m + bQ_n$ for all constants a and b. The argument used in Theorem 2.4 (or rather first used in Corollary 2.2) is thus no longer valid. And it cannot be successfully amended.

$$K(t) \cong r_1 J(t) + r_2 J^2(t) + \cdots . \tag{3.26}$$

Similarly,

$$\begin{cases} K^i(t) \cong r_{ii} J^i(t) + r_{i,i+1} J^{i+1}(t) + \cdots & (i = 1, 2, \cdots), \\ \displaystyle\sum_{n=i}^{\infty} r_{in} t^n \cong \left(\sum_{n=1}^{\infty} r_n t^n \right)^i . \end{cases} \tag{3.27}$$

Recalling that $J[P_n] = P_{n-1}$, we find that

$$\begin{cases} K[P_n] = r_1 P_{n-1} + r_2 P_{n-2} + \cdots , \\ K^2[P_n] = \qquad\qquad r_{22} P_{n-2} + r_{23} P_{n-3} + \cdots , \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ K^i[P_n] = \qquad\qquad\qquad r_{ii} P_{n-i} + r_{i,i+1} P_{n-i-1} + \cdots . \end{cases} \tag{3.28}$$

And on using these relations and also (3.21) in (3.23), we obtain the further relations

$$\begin{cases} nP_n = \displaystyle\sum_{j=1}^{n} S_{1j} \{ r_{jj} P_{n-j} + r_{j,j+1} P_{n-j-1} + \cdots + r_{jn} P_0 \} & (n = 1, 2, \cdots); \\ n(n-1)P_n = \displaystyle\sum_{j=2}^{n} S_{2j} \{ r_{jj} P_{n-j} + \cdots + r_{jn} P_0 \} & (n = 2, 3, \cdots); \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ n(n-1) \cdots (n-K+1) P_n & \\ \qquad = \displaystyle\sum_{j=k}^{n} S_{kj} \{ r_{jj} P_{n-j} + \cdots + r_{jn} P_0 \} & (n = k, k+1, \cdots). \end{cases} \tag{3.29}$$

Equations (3.29) enable us to determine the $S_{ij}(x)$'s in terms of the set P.

Let us define $\Theta(t)$ by

$$\Theta(t) \cong r_1 t + r_2 t^2 + \cdots . \tag{3.30}$$

Comparison with (3.26) shows that

$$\Theta(J(t)) \cong K(t). \tag{3.31}$$

Suppose the first relation of (3.29) is multiplied by t^n and the result summed from $n = 1$ to $n = \infty$. The left side becomes $\sum nP_n t^n$, while the right side is seen to reduce to

$$\left(\sum_0^{\infty} P_n t^n \right) \left\{ S_{11} \sum_1^{\infty} r_{1n} t^n + S_{12} \sum_2^{\infty} r_{2n} t^n + \cdots \right\}.$$

From the definition of r_{ij}, the series in the braces represent $\Theta(t)$, $\Theta^2(t)$, $\cdots$. That is,

$$\sum_{n=1}^{\infty} nP_n t^n \cong \left(\sum_0^{\infty} P_n t^n \right) \{ S_{11} \Theta + S_{12} \Theta^2 + \cdots \}.$$

Now $\sum P_n t^n$ is given by (2.7). We thus have

$$t\frac{\partial}{\partial t}\{A(t)e^{xH(t)}\} \cong A(t)e^{xH(t)}\{S_{11}(x)\Theta(t) + S_{12}(x)\Theta^2(t) + \cdots\}. \tag{3.32}$$

In similar manner we obtain from the second relation of (3.29),

$$t\frac{\partial^2}{\partial t^2}\{Ae^{xH}\} \cong Ae^{xH}\{S_{22}\Theta^2 + S_{23}\Theta^3 + \cdots\}; \tag{3.33}$$

and from the general relation of (3.29),

$$t^k\frac{\partial^k}{\partial t^k}\{Ae^{xH}\} \cong Ae^{xH}\{S_{kk}\Theta^k + S_{k,k+1}\Theta^{k+1} + \cdots\}. \tag{3.34}$$

These relations permit us to establish

THEOREM 3.6. *Let P be of type zero, with operator J and determining function A. In order that P satisfy a finite order equation of form (3.13) it is necessary and sufficient that constants $\alpha_1, \cdots, \alpha_r$ exist, not all zero, such that the function*

$$Q(t, x) \equiv \frac{e^{-xH(t)}}{A(t)}\left[\alpha_1 t\frac{\partial}{\partial t}\{Ae^{xH}\} + \cdots + \alpha_r t\frac{\partial^r}{\partial t^r}\{Ae^{xH}\}\right], \tag{3.35}$$

when expressed as a power series in $\Theta(t)$, reduces to a polynomial in $\Theta(t)$.

Q can be written as a power series in t, and can therefore (formally) be expressed as a series of powers of $\Theta(t)$. More precisely, from (3.32) to (3.34) we have

$$\begin{aligned} Q(t, x) \cong \{\alpha_1 S_{11}\}\Theta(t) + \cdots + \{\alpha_1 S_{1r} + \cdots + \alpha_r S_{rr}\}\Theta^r(t) \\ + \sum_{n=r+1}^{\infty}\{\alpha_1 S_{1n} + \cdots + \alpha_r S_{rn}\}\Theta^n(t). \end{aligned} \tag{3.36}$$

Suppose P satisfies a finite order equation, so that conditions (3.25) hold for some r. Then $Q(t, x)$, as seen by (3.36), reduces to a polynomial in $\Theta(t)$. The necessity of Theorem 3.6 is thus proved. Conversely, suppose that for some r $Q(t, x)$ is a polynomial in $\Theta(t)$. We wish to show that set P, corresponding to the A and J in terms of which Q is defined, satisfies a finite order equation. We know that P satisfies (2.17). Using operator T of (2.17), we form the equation

$$L[y] \equiv \sum_{k=1}^{r}\alpha_k T_k[y] = \lambda y, \tag{3.37}$$

which is also satisfied by P. It is this equation that we shall prove is of finite order.

If $L[y]$ is recast in terms of $K[y]$, so that it is of form (3.13), the coefficients $L_n(x)$ are given by (3.24), with $\alpha_n = 0$ for $n > r$. Now from (3.36), since Q is a polynomial in $\Theta(t)$, there is an integer s such that

$$\alpha_1 S_{1n} + \cdots + \alpha_r S_{rn} = 0, \qquad n > s.$$

Hence $L_n(x) = 0$ for $n > s$. This establishes the sufficiency.

The condition of Theorem 3.6 is more serviceable than is that of Lemma 3.2, since the function Q is determined directly from A and H. We can rid ourselves of the function $\Theta(t)$ on making use of (3.31). It gives us

COROLLARY 3.3. *The zero type set P satisfies a finite order equation if and only if constants $\alpha_1, \cdots, \alpha_r$ (not all zero) exist so that $Q(J(t), x)$ is a polynomial in the function $K(t)$.*

Whenever the choice $r = 1$ is permissible, the condition of Theorem 3.6 (or of Corollary 3.3) is seen to reduce to the conditions (3.2), (3.3) already met.

COROLLARY 3.4. *If P satisfies an r-th order equation* (3.13), *then for this equation Q is given by*

$$Q(t, x) = \sum_{j=1}^{r} L_j(x)\Theta^j(t). \tag{3.38}$$

This follows from (3.36).

Corollary 3.4 enables us to show that neither the Legendre set $\{X_n(x)\}$ nor any set $\{c_n X_n\}$ is of type zero. The Legendre polynomials are given by

$$(1 - 2tx + t^2)^{-\frac{1}{2}} = \sum_0^\infty X_n(x)t^n.$$

If $\{X_n\}$ is of type zero, then the left member is of the form $\exp\{xH(t)\}$. This is readily seen to be impossible.

Now suppose $\{P_n = c_n X_n\}$ $(c_n \neq 0)$ is of type zero. X_n, and therefore P_n, satisfies the finite order equation

$$(1 - x^2)y'' - 2xy' = \lambda y$$

with $\lambda = -n(n + 1)$ for $y = P_n$. Here the operator $K[y]$ is merely $y'(x)$, so that $\Theta(t) = H(t)$. Also, $L_1 = -2x$, $L_2 = 1 - x^2$. Hence, from the corollary,

$$Q(t, x) = L_1 H + L_2 H^2 = -2xH + (1 - x^2)H^2.$$

If we equate coefficients of like powers of x on both sides, we get from the x^2 terms: $t^2 H'^2 = H^2$, so that $H = ct$; and on using this result in the equation obtained from the x terms, we find that $A(t) \equiv$ constant. Finally, the constant terms tell us that $H \equiv 0$ so that $c = 0$. Hence, $\sum P_n t^n$ has for sum a constant. This contradiction shows that $\{P_n = c_n X_n\}$ is not of type zero.

4. Zero type sets that are Tchebycheff sets.

The Hermite polynomials are Appell polynomials, and are thus of zero type. They are also Tchebycheff orthogonal polynomials.[18] Another orthogonal set of zero type is the Laguerre

[18] The definition of $H_n(x)$ in Example 3 of §2 requires modification in order to satisfy the condition $H_n'(x) = H_{n-1}(x)$ for an Appell set. But such alteration consists only in multiplying each H_n by a suitable non-zero constant c_n. This being done, it is known that the Hermite set is essentially the only Tchebycheff set that is also an Appell set.

set (cf. (2.10)). This suggests the problem of determining all zero type sets that are orthogonal.

J. Meixner[19] has treated this problem by the use of the Laplace transformation, taking (essentially) the relation (2.7) as the definition of the polynomials under consideration. It is possible to give a quite different treatment by means of the known properties of zero type sets, and this we do here.

As a characterization of an orthogonal set $\{Q_n\}$ we take the relation[20]

$$Q_n(x) = (x + \lambda_n)Q_{n-1}(x) + \mu_n Q_{n-2}(x) \qquad (n = 1, 2, \cdots), \tag{4.1}$$

λ_n, μ_n being real constants with $\mu_n \neq 0$, $n > 1$. If $\{Q_n\}$ is an orthogonal set, so is $\{c_n Q_n\}$, $c_n \neq 0$ (although the multipliers c_n can spoil normality if Q_n has this latter property). We shall therefore set the problem as follows: *For what sets* $\{Q_n\}$ *satisfying* (4.1) *do there exist non-zero constants* c_n *such that*

$$P_n(x) = c_n Q_n(x) \qquad (n = 0, 1, \cdots) \tag{4.2}$$

is a set of type[21] *zero?*

Suppose that $\{P_n\}$ of (4.2) is of type zero. From (4.1) we obtain an expression for $nP_n(x)$. Comparing this with the value of $nP_n(x)$ as given by (2.20), we obtain (on equating coefficients of like powers of x):

$$n!c_n = c_0 q_{11}^n, \qquad q_{11}^2 \lambda_n = q_{10}q_{21} + (n-1)q_{21}, \tag{4.3}$$

$$q_{11}^4 \mu_{n+1} = n\{q_{20}q_{11}^2 - q_{10}q_{11}q_{21} - (q_{21}^2 - q_{11}q_{31})(n-1)\}. \tag{4.4}$$

That is, λ_n is at most linear and μ_n at most quadratic, in n, and μ_n has a factor $(n - 1)$.

[19] *Orthogonale Polynomsysteme mit einer besonderern Gestalt der erzeugenden Funktion*, Journal of London Math. Soc., vol. 9 (1934), pp. 6–13.

[20] As justification we observe first that every set orthogonal according to the classical definition satisfies a relation of form (4.1); and secondly, that Shohat has shown that a necessary and sufficient condition that a set $\{Q_n\}$ (normalized so that the x^n term has a coefficient unity) be orthogonal with respect to a weight function $\psi(x)$ of bounded variation in $(-\infty, +\infty)$ is that $\{Q_n\}$ satisfy (4.1) with $\mu_n \neq 0$ for all $n > 1$. (J. Shohat, Comptes Rendus, vol. 207(1938), pp. 556–558.)

We note that in the Shohat definition of orthogonality it is tacitly assumed that no member of an orthogonal set is orthogonal to itself (relative to the given weight function ψ). It is easy to show that if an "orthogonal" set satisfies (4.1) with $\mu_n = 0$ for some $n > 1$, then at least one polynomial of the set is self-orthogonal. Thus, for example, the set $\{x^n\}$ is "orthogonal" for the following choice of ψ:

$$\psi(x) = \begin{cases} 1, & x = 0; \\ 0, & x \neq 0. \end{cases}$$

This set also satisfies (4.1) with $\lambda_n = \mu_n = 0$ for all n. This apparent contradiction to the theorem of Shohat is resolved when we note that x^n is orthogonal to itself for every $n > 0$.

A colleague, H. L. Krall, has made the further observation that every set P for which $P_n(0) = 0$ $(n > 0)$ is "orthogonal" relative to the same function ψ above.

[21] We shall also assume for convenience that $P_0(x) = 1$. This is no essential restriction since it means only that all the polynomials $P_n(x)$ are multiplied by *one and the same* non-zero constant. The property of being of zero type is thus unaltered.

This condition is also sufficient. For suppose λ_n, μ_n have this form, and let $\{Q_n\}$ be defined by (4.1). Let $\alpha \neq 0$ be any number and set

$$c_n = \alpha^n/n! \qquad (n = 0, 1, \cdots). \tag{4.5}$$

Now define $\{P_n\}$ by (4.2). We are to show that P is of type zero. From (4.1) and (4.2) follows a relation of form

$$nP_n = (\alpha x + \beta + n\gamma)P_{n-1} + (\delta + n\epsilon)P_{n-2}, \tag{4.6}$$

$\alpha \neq 0$, $\delta + n\epsilon \neq 0$ $(n > 1)$. From this we obtain xP_{n-1} as a linear combination of P_n, P_{n-1}, P_{n-2}. It is now a straightforward matter to show that constants q_{k0}, q_{k1} exist so that

$$T_n \equiv (q_{10} + xq_{11})P_{n-1} + (q_{20} + xq_{21})P_{n-2} + \cdots$$

is identically equal to nP_n. They are, in fact, determined by the relations

$$q_{k+2,1} = \gamma q_{k+1,1} + \epsilon q_{k1}, \tag{4.7}$$

$$q_{k+1,0} = \frac{1}{\alpha}\{q_{k1}(\delta - (k-1)\epsilon) + q_{k+1,1}(\beta - \gamma k) + (k+1)q_{k+2,1}\}. \tag{4.8}$$

Thus (2.20) holds, and P is of type zero. That is, we can state

THEOREM 4.1. *A necessary and sufficient condition that an orthogonal set* $\{Q_n\}$, *given by* (4.1), *be such that* $P_n = c_nQ_n$ *is of type zero for some choice of* $c_n \neq 0$ *is that* λ_n, μ_n *have the form*

$$\lambda_n = a + bn, \qquad \mu_n = (n-1)(c + dn), \tag{4.9}$$

with $c + dn \neq 0$ *for* $n > 1$.

As it stands this criterion does not reveal the sets P that are both orthogonal and of type zero. We therefore examine the problem more closely. Relation (4.8) shows that $\{q_{k0}\}$ is determined when $\{q_{k1}\}$ is known. Let us then turn to the recurrence relation (4.7). The characteristic equation is

$$u^2 - \gamma u - \epsilon = 0. \tag{4.10}$$

Case I. $\gamma^2 + 4\epsilon = 0$. Using the initial conditions $q_{11} = \alpha$, $q_{21} = \alpha\gamma$, we obtain

$$q_{k1} = \alpha k(\tfrac{1}{2}\gamma)^{k-1}, \qquad q_{k+1,0} = (\tfrac{1}{2}\gamma)^{k-1}\{\tfrac{1}{2}(\beta\gamma + 2\delta)k + \tfrac{1}{2}\gamma(\beta + \gamma)\}. \tag{4.11}$$

Then from (2.12) and (2.13), we get[22]

[22] The presence of the parameter μ removes the earlier condition that $P_0(x) = 1$. It should also be noticed that we must have $\gamma \neq 0$ in (4.18). The case $\gamma = 0$ is special. For if $\gamma = 0$, then $\epsilon = 0$, and

$$q_{11} = \alpha, \quad q_{k1} = 0\ (k > 1); \quad q_{10} = \beta, \quad q_{20} = \delta, \quad q_{k0} = 0\ (k > 2).$$

Hence

$$H(t) = \alpha t, \quad J(t) = t/\alpha, \quad A(t) = \mu\cdot\exp\{\beta t + \tfrac{1}{2}\delta t^2\},$$

where $\delta \neq 0$ in order that the condition $\mu_n \neq 0$ be fulfilled.

$$
(4.12)\qquad \begin{aligned} H(t) &= \frac{2\alpha t}{2 - \gamma t}, \qquad J(t) = \frac{2t}{2\alpha + \gamma t}, \\ A(t) &= \mu\left(\frac{2 - \gamma t}{2}\right)^{\omega} \cdot \exp\left\{\frac{4(\beta\gamma + 2\delta)}{\gamma^2(2 - \gamma t)}\right\}, \end{aligned}
$$

where

$$
\omega \equiv \frac{-2}{\gamma^2}(\gamma^2 - 2\delta) \neq \text{a non-negative integer}
$$

(in order that $\mu_n \neq 0$).

Case II. $\gamma^2 + 4\epsilon \neq 0$. Let u_1, u_2 be the roots of (4.10). Then

$$
(4.13)\qquad q_{k1} = \alpha(\gamma^2 + 4\epsilon)^{-\frac{1}{2}}\{u_1^k - u_2^k\}, \qquad q_{k+1,0} = (\gamma^2 + 4\epsilon)^{-\frac{1}{2}}\{\lambda u_1^k - \sigma u_2^k\},
$$

where λ, σ are constants whose values are readily obtained. Consequently

$$
(4.14)\qquad H(t) = \alpha \int_0^t \frac{dt}{1 - \gamma t - \epsilon t^2}.
$$

Case II_1. $\epsilon = 0$ (so that $\gamma \neq 0$). Then

$$
(4.15)\qquad \left\{ \begin{aligned} H(t) &= \frac{-\alpha}{\gamma}\log(1 - \gamma t), \qquad J(t) = \frac{1}{\gamma}\left(1 - \exp\left\{-\frac{\gamma t}{\alpha}\right\}\right), \\ A(t) &= \mu(1 - \gamma t)^{-\theta} \cdot \exp\left\{-\frac{\delta t}{\gamma}\right\}, \end{aligned} \right.
$$

where

$$
\theta = \frac{1}{\gamma^2}(\beta\gamma + \gamma^2 + \delta),
$$

and where $\delta \neq 0$ in order that $\mu_n \neq 0$.

Case II_2. $\epsilon \neq 0$. Let $r_i = 1/u_i$. Then

$$
(4.16)\qquad H(t) = \frac{1}{\rho}\log\left\{\frac{r_1(t - r_2)}{r_2(t - r_1)}\right\}, \quad J(t) = \frac{e^{\rho t} - 1}{u_1 e^{\rho t} - u_2}, \quad A(t) = \mu \cdot \frac{(1 - u_2 t)^{h_2}}{(1 - u_1 t)^{h_1}},
$$

where

$$
(4.17)\qquad \rho = \frac{1}{\alpha}(\gamma^2 + 4\epsilon)^{\frac{1}{2}}, \qquad h_i = (\gamma^2 + 4\epsilon)^{-\frac{1}{2}}\,\frac{u_i(\beta + \gamma) + (\delta + 2\epsilon)}{u_i}.
$$

It is to be noted that in all cases H and J do not involve the parameters β, δ, μ and A does not involve α. It follows that *all sets satisfying a relation of form* (4.6) *and having the same* α, γ, ϵ *correspond to the same operator* J; *and all sets satisfying* (4.6) *with the same* β, γ, δ, ϵ *have the same determining*[23] *function* A.

The many relations obtained for H, J, A involve the original parameters $\alpha, \cdots, \epsilon$ and μ, sometimes in complicated manner. This suggests the possi-

[23] At least to within a constant multiplier (because of the presence of μ).

bility of simplifying by introducing independent combinations of the original parameters as new parameters. In fact, the following relations summarize the various cases. They are, in order: Cases I, I special,[24] II_1, II_2; that is,

$$(4.18)\qquad H(t) = \frac{at}{1 - bt}, \qquad J = \frac{t}{a + bt}, \qquad A = \mu(1 - bt)^c \cdot \exp\left\{\frac{d}{1 - bt}\right\},$$

where a, b, c, d, μ are arbitrary, but $abc\mu \neq 0$.

$$(4.19)\qquad H = at, \qquad J = \frac{t}{a}, \qquad A = \mu \exp\{bt + ct^2\},$$

a, b, c, μ arbitrary, but $ac\mu \neq 0$.

$$(4.20)\qquad H = a \log(1 - bt), \qquad J = \frac{1}{b}\{1 - e^{t/a}\}, \qquad A = \mu e^{ct} \cdot (1 - bt)^d,$$

a, b, c, d, μ arbitrary, but $abc\mu \neq 0$.

$$(4.21)\quad H = \frac{1}{a}\log\left\{\frac{b(t - c)}{c(t - b)}\right\}, \quad J = bc\left(\frac{e^{at} - 1}{ce^{at} - b}\right), \quad A = \mu\left(1 - \frac{t}{c}\right)^{d_1} \cdot \left(1 - \frac{t}{b}\right)^{d_2},$$

a, b, c, d_1, d_2, μ arbitrary, but $abc\mu \neq 0$ and[25] $b \neq c$.

We therefore have

THEOREM 4.2. *Let P be of zero type, with operator J and determining function A, so that* (2.7) *holds. A necessary and sufficient condition that P be an orthogonal set is that J, H, A satisfy one of the conditions* (4.18) *to* (4.21).

COROLLARY 4.1. *According to the case, the function Ae^{xH} of* (2.7) *assumes the form:*[26]

$$(4.22)\qquad Ae^{xH} = \mu(1 - bt)^c \cdot \exp\left\{\frac{d + atx}{1 - bt}\right\} \qquad (abc\mu \neq 0),$$

$$(4.23)\qquad Ae^{xH} = \mu \cdot \exp\{t(b + ax) + ct^2\} \qquad (ac\mu \neq 0),$$

$$(4.24)\qquad Ae^{xH} = \mu e^{ct} \cdot (1 - bt)^{d + ax} \qquad (abc\mu \neq 0),$$

$$(4.25)\qquad Ae^{xH} = \mu\left(1 - \frac{t}{c}\right)^{d_1 + x/a} \cdot \left(1 - \frac{t}{b}\right)^{d_2 - x/a} \qquad (abc\mu \neq 0,\ b \neq c).$$

The Laguerre set is a particular case of (4.22) and the Hermite set of (4.23). If it were permissible to choose $c = 0$ in (4.23) and (4.24) we would have as particular cases, respectively, $\{x^n/n!\}$ and the Newton set. Hence, these two sets, while not Tchebycheff sets, are nevertheless limiting sets of Tchebycheff sets.

If we form the functions $\{uA'(u)/A(u)\}$, $\{uH'(u)\}$, evaluated for $u = J(t)$, we find in the respective cases that

[24] Case I special refers to an earlier footnote under Case I.

[25] Also, the condition $\delta + \epsilon n \neq 0$ $(n > 1)$ is to be translated in terms of the present parameters.

[26] (4.25) is subject to the restriction mentioned in the preceding footnote.

$$(4.26)\quad \left\{\frac{uA'}{A}\right\} = \frac{bt}{a^2}[a(d-c)+bdt], \qquad \{uH'\} = \frac{t}{a}(a+bt);$$

$$(4.27)\quad \left\{\frac{uA'}{A}\right\} = \frac{bt}{a} + \frac{2ct^2}{a^2}, \qquad \{uH'\} = t;$$

$$(4.28)\quad \left\{\frac{uA'}{A}\right\} = (1-e^{t/a})(ce^{t/a}-bd) \div be^{t/a}, \qquad \{uH'\} = a(e^{t/a}-1)e^{-t/a};$$

$$(4.29)\quad \left\{\frac{uA'}{A}\right\} = \frac{bc(1-e^{at})}{(c-b)}\left[\frac{d_1}{ce^{at}} + \frac{d_2}{b}\right], \qquad \{uH'\} = \frac{(e^{at}-1)(ce^{at}-b)}{a(c-b)e^{at}}.$$

Since in (4.26) and (4.27) the expressions are polynomials, it follows from Theorem 2.3 that the sets P of the first two cases satisfy the respective finite order equations

$$(4.30)\quad M[y(x)] \equiv \left\{\frac{b}{a}(d-c)+x\right\}y'(x) + \left\{\frac{b^2d}{a^2}+\frac{b}{a}x\right\}y''(x) = \lambda y(x),$$

$$(4.31)\quad M[y(x)] \equiv \left\{\frac{b}{a}+x\right\}y'(x) + \frac{2c}{a^2}y''(x) = \lambda y(x),$$

where $\lambda = n$ for $y = P_n$. The sets for the last two cases clearly do not satisfy a finite order equation of form (2.14).

5. **Sets of higher type.** Although the present paper has as its main purpose the treatment of zero type sets, we propose in this section to indicate some extensions to sets of higher type. The definition of higher type depends on what characterization of zero type sets one wishes to generalize. We have given one definition in §1. This we shall call A-type. Thus: *A set P is of A-type k if in* (1.12) *the maximum degree of the coefficients $L_n(x)$ is k. If the L_n's are of unbounded degree, P is of infinite A-type.*

Let P be an arbitrary set. There exists a unique sequence of formal power series $\{M_n(t)\}$ of form

$$(5.1)\qquad M_n(t) \cong m_{nn}t^n + m_{n,n+1}t^{n+1} + \cdots \qquad (m_{nn} \neq 0)$$

such that[27]

$$(5.2)\qquad e^{tx} \cong \sum_{n=0}^{\infty} P_n(x)M_n(t).$$

Definition. We shall term the set of series (or functions) M: $\{M_n(t)\}$ *the E-associate of set P.*

Corollary 5.1. *Let P be a set and M its E-associate. A necessary and sufficient condition that P be of type zero is that formal series*

[27] For our purpose it is a matter of indifference whether or not the series symbolized by $M_n(t)$ converge. (5.1) is regarded as a "carrier" for the coefficients m_{nk}, and (5.2) is merely a concise way of writing infinitely many linear equations in these coefficients.

$$(5.3)\qquad A(t) \cong \sum_0^\infty a_n t^n \quad (a_0 \neq 0), \qquad J(t) \cong \sum_1^\infty c_n t^n \quad (c_1 \neq 0)$$

exist so that

$$(5.4)\qquad M_n(t) = \frac{[J(t)]^n}{A(J(t))} = \left\{\frac{u^n}{A(u)}\right\} \qquad (u = J(t)).$$

For, if P is of type zero, (2.7) holds. And from (5.2), on setting $t = H(u)$,

$$A(u)e^{xH(u)} \cong \sum P_n(x)M_n(H(u))A(u),$$

so that

$$A(u)M_n(H(u)) = u^n.$$

(5.4) follows on inverting: $u = J(t)$.

Conversely, suppose (5.3) and (5.4) are satisfied. Define the set Q to be of type zero, corresponding to operator J and determining function A, and let M^* be its E-associate. Then from the half already established, M_n^* has the value given by (5.4). That is, $M_n^*(t) = M_n(t)$ for all n. But just as M is uniquely determined from knowledge of P and (5.2), so is P uniquely defined by (5.2) when M is given. Hence, Q and P are identical, and P is of type zero.

We now characterize sets of A-type k. Let P be such a set. Then, as we know,

$$(5.5)\qquad L[P_n] = P_{n-1},$$

where

$$(5.6)\qquad L[y(x)] \equiv J_0[y] + xJ_1[y] + \cdots + x^kJ_k[y],$$

$J_0, \cdots, J_k$ being linear differential operators with constant coefficients such that

$$(5.7)\qquad J_i(t) \cong a_{i,i+1}t^{i+1} + a_{i,i+2}t^{i+2} + \cdots, \qquad 0 \leqq i \leqq k.$$

Also, in order to insure that L carries every polynomial into another of degree one less (since $L[P_n] = P_{n-1}$), we have the further condition

$$(5.8)\qquad \xi_n = a_{01} + na_{12} + n(n-1)a_{23} + \cdots + n \cdots (n-k+1)a_{k,k+1} \neq 0 \qquad (n = 0, 1, \cdots).$$

From (5.6) and (5.2) we have

$$\text{(a)}\qquad \begin{aligned} L[e^{tx}] &\cong \{J_0(t) + xJ_1(t) + \cdots + x^kJ_k(t)\}e^{tx} \\ &\cong \sum_0^\infty L[P_n]M_n(t) \cong \sum_0^\infty P_n(x)M_{n+1}(t). \end{aligned}$$

But also, on differentiating (5.2) k times in t, we get

$$\text{(b)} \qquad \{J_0 + \cdots + x^k J_k\}e^{tx} \cong \sum_0^\infty P_n(x)\{J_0 M_n + J_1 M_n' + \cdots + J_k M_n^{(k)}\}.$$

Again, in (a) and (b) both the coefficients of $P_n(x)$ are power series beginning with a term in t^{n+1} (cf. relation (5.8)). These coefficients must therefore be identical. That is,

$$\text{(5.9)} \quad M_{n+1}(t) = J_0(t)M_n(t) + J_1(t)M_n'(t) + \cdots + J_k(t)M_n^{(k)}(t) \quad (n = 0, 1, \cdots).$$

This proves the necessary part of

THEOREM 5.1. *Let P be a set and M its E-associate. A necessary and sufficient condition that P be of finite* A*-type is that for some k there exist formal power series* (5.7) *satisfying condition* (5.8) *such that M satisfies* (5.9). *Moreover, if $J_k(t) \not\equiv 0$, P is of* A*-type k.*

The sufficiency is established as follows: (5.7) and (5.8) determine an operator L of form (5.6). The first part of (a) holds to give

$$\text{(c)} \qquad L[e^{tx}] \cong \sum_0^\infty L[P_n]M_n \cong \sum_0^\infty L[P_{n+1}]M_{n+1},$$

while (b) continues to hold. On using (5.9), (b) reduces to

$$\text{(d)} \qquad L[e^{tx}] \cong \sum_0^\infty P_n M_{n+1},$$

so that

$$\text{(e)} \qquad \sum_0^\infty \{L[P_{n+1}] - P_n\}M_{n+1}(t) = 0.$$

This formal identity means that if we rearrange in a power series in t, all coefficients must vanish. Recalling the form (5.1) of the functions M_n, we see that all the braces likewise vanish. That is, $L[P_n] = P_{n-1}$, $n \geqq 1$. That $L[P_0] = 0$ is immediate. Hence P is of A-type $\leqq k$. The statement in the theorem concerning the precise type number is evident.

If $k = 0$, (5.9) reduces to $M_{n+1} = JM_n$ (dropping the subscript 0); i.e., $M_n(t) = M_0(t)[J(t)]^n$. This coincides with the previously found condition (5.4) if $A(t)$ is chosen so that $M_0(t)A(J(t)) = 1$.

It has already been observed that if M is any sequence of power series of form (5.1) (which includes the condition $m_{nn} \neq 0$), then (5.2) uniquely defines a set P for which M is the E-associate. To say that M is the E-associate of some set P is then equivalent to saying that M is of form (5.1).

From Theorem 5.1 we can prove

THEOREM 5.2. *Let M be of form* (5.1), *so that M is the E-associate of a set P. A necessary and sufficient condition that P be of finite* A*-type is that k exist so that the following* $(k + 1)$ *ratios of determinants are independent of n:*

$$\Delta_j(t) = \begin{vmatrix} M_n & \cdots & M_n^{(j-1)} & M_{n+1} & M_n^{(j+1)} & \cdots & M_n^{(k)} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ M_{n+k} & \cdots & M_{n+k}^{(j-1)} & M_{n+k+1} & M_{n+k}^{(j+1)} & \cdots & M_{n+k}^{(k)} \end{vmatrix} \tag{5.10}$$
$$\div \begin{vmatrix} M_n & M_n' & \cdots & M_n^{(k)} \\ \cdots & \cdots & \cdots & \cdots \\ M_{n+k} & M_{n+k}' & \cdots & M_{n+k}^{(k)} \end{vmatrix} \qquad (j = 0, 1, \cdots, k).$$

And if $\Delta_k(t) \not\equiv 0$, *the* A-*type is precisely* k.

First, suppose P is of A-type k. If in (5.9) we replace n by n, $n + 1$, $\cdots$, $n + k$ respectively, we obtain $k + 1$ equations for $J_0, \cdots, J_k$ whose solution is given by the determinant ratios[28] in (5.10); i.e., $J_j(t) = \Delta_j(t)$. Thus the condition is necessary. Conversely, suppose (5.10) holds, the Δ_j's being independent of n. Then (5.9) is satisfied by the choice $J_j = \Delta_j$. The sufficiency will now follow from Theorem 5.1 if we show that J_j satisfies conditions (5.7) and (5.8).

Suppose the series (5.1) is substituted into (5.10). It is found that the numerator and denominator have as lowest terms, respectively,

$$\alpha_{jn} t^{nk+j+1} \cdot \prod_{i=n}^{n+k} m_{ii}, \qquad \beta_n t^{nk} \cdot \prod m_{ii},$$

where

$$\beta_n = \begin{vmatrix} 1 & n & \cdots & n(n-1) \cdots (n-k+1) \\ 1 & (n+1) & \cdots & (n+1) \cdots (n-k+2) \\ \cdots & \cdots & \cdots & \cdots \\ 1 & (n+k) & \cdots & (n+k) \cdots (n+1) \end{vmatrix},$$

and where α_{jn} is obtained by replacing the $(j + 1)$-th column of β_n by the elements $m_{i+n,i+n} \div m_{i+n-1,i+n-1}$ $(i = 1, 2, \cdots, k + 1)$. Hence J_j cannot begin with a term of degree less than t^{j+1} if we show that $\beta_n \neq 0$. If in β_n we subtract each row from the one following, we obtain a new determinant (of value β_n), which is $k!$ times β_n', where β_n' is obtained from β_n by changing k to $k - 1$. Working down to $k = 1$ we get $\beta_n = k!(k-1)! \cdots 2!\,1!$, and this is not equal to zero.

Thus, condition (5.7) holds. There remains to establish (5.8). Since $\Delta_j = J_j$ begins with the term $t^{j+1}\alpha_{jn} \div \beta_n$ (or a term of even higher degree), we see on comparing with (5.7) that $\alpha_{j,j+1} = \alpha_{jn} \div \beta_n$. Hence $\alpha_{j,j+1}$ $(j = 0, 1, \cdots, k)$ is the solution of a system of $(k + 1)$ linear non-homogeneous equations, the matrix of whose coefficients is given by the elements of the determinant β_n, and whose non-homogeneous terms are the quantities $m_{i+n,i+n} \div m_{i+n-1,i+n-1}$. The first of these equations is

$$a_{01} + na_{12} + \cdots + n(n-1) \cdots (n-k+1) a_{k,k+1} = \frac{m_{n+1,n+1}}{m_{nn}}.$$

[28] That the denominator does not vanish identically is demonstrated later in the proof, where it is shown that $\beta_n \neq 0$.

The left side is the quantity ξ_n of (5.8). As $m_{ii} \neq 0$ for all i, so is $\xi_n \neq 0$ for all n. Thus (5.8) holds and the proof is complete.

We come now to a second definition of type.

LEMMA 5.1. *To each set P corresponds a unique sequence of polynomials $\{T_n(x)\}$, with $T_n(x)$ of degree not exceeding n, such that*[29]

$$P_n' = T_0P_{n-1} + T_1P_{n-2} + \cdots + T_{n-1}P_0 \qquad (n = 1, 2, \cdots). \tag{5.11}$$

This is seen if we set $n = 1, 2, \cdots$ successively. It is to be observed that the T_n's do not determine the P_n's uniquely. In fact, there are infinitely many sets satisfying (5.11) for a given sequence $\{T_n\}$.

DEFINITION. *A set P is of* B-*type k if in* (5.11) *the maximum degree of the polynomials $T_n(x)$ is k. Otherwise, P is of infinite* B-*type.*

If P is of B-type zero, the T_n's are constants, so that (5.11) reduces to (2.21). That is, P is of A-type zero. The converse is also true:

COROLLARY 5.2. *A set P is of* B-*type zero if and only if it is of* A-*type zero.*

Let

$$H(x, t) \cong \sum_0^\infty P_n(x)t^n, \tag{5.12}$$

$$T(x, t) \cong \sum_0^\infty T_n(x)t^n. \tag{5.13}$$

Then (5.11) is seen to be equivalent to the relation

$$tHT = \frac{\partial H}{\partial x},$$

which, when solved for H, gives us

$$H(x, t) = A(t) \exp\left\{t \int_0^x T(x, t)\, dx\right\}, \tag{5.14}$$

where it is to be understood here and later that $A(t)$ is an arbitrary power series beginning with a (non-zero) constant term. Conversely, if T is any series (5.13) where T_n is a polynomial of degree not exceeding n, then $H(x, t)$ as defined by (5.14) is such that (5.11) holds.[30]

If $T(x, t)$ is written as a power series in x rather than t, (5.14) assumes the form

$$H(x, t) \cong A(t) \exp\{xH_1(t) + x^2H_2(t) + \cdots\}, \tag{5.15}$$

[29] This is an extension of Theorem 2.6.

[30] The T_n's are not completely arbitrary. For P to be a set, it is necessary that P_n be of degree exactly n. This reflects itself in the non-vanishing for $n = 1, 2, \cdots$ of certain polynomials in $t_{00}, t_{11}, t_{22}, \cdots$, where t_{ii} is the coefficient of x^i in $T_i(x)$. These conditions can be obtained from (5.11) by demanding that the right member be of degree exactly $(n - 1)$.

where the H_i are power series in t, H_i beginning with a term in t^i or possibly higher. (H_1 definitely begins with a term in t.)

For P to be of B-type k, $T(x, t)$ must be a polynomial in x of degree k. This is necessary and sufficient. On integrating, we get a polynomial of degree $k + 1$ in x. Hence we have

THEOREM 5.3. *A necessary and sufficient condition that a set P be of* B*-type k is that it be given by* (5.12) *where H is of the form*

$$H(x, t) = A(t)\cdot \exp\{xH_1(t) + \cdots + x^{k+1}H_{k+1}(t)\}, \tag{5.16}$$

the $H_i(t)$ being of form[31]

$$H_i(t) \cong h_{ii}t^i + h_{i,i+1}t^{i+1} + \cdots \qquad (h_{11} \neq 0). \tag{5.17}$$

Given a set P, there exists a unique sequence of polynomials $\{U_n(x)\}$, U_n of degree not exceeding n, such that

$$nP_n = U_1P_{n-1} + \cdots + U_nP_0 \qquad (n = 1, 2, \cdots). \tag{5.18}$$

(The U_n's are determined successively if we set $n = 1, 2, \cdots$.)

DEFINITION. *P is of* C*-type k if the maximum degree of the U_n's is $(k + 1)$.*

If we set

$$U(x, t) \cong \sum_0^\infty U_{n+1}(x)t^n, \tag{5.19}$$

then from (5.12) and (5.18),

$$HU = \frac{\partial H}{\partial t},$$

so that

$$H(x, t) = c\cdot\exp\left\{\int_0^t U(x, t)\,dt\right\}. \tag{5.20}$$

Here c is an arbitrary (non-zero) constant. Comparing (5.14) and (5.20), we see that

$$\log c + \int_0^t U(x, t)\,dt = \log A(t) + t\int_0^x T(x, t)\,dx, \tag{5.21}$$

so that

$$\begin{cases} U(x, t) = \dfrac{A'}{A} + \displaystyle\int_0^x \frac{\partial}{\partial t}(tT)\,dx, \\ tT(x, t) = \displaystyle\int_0^t \frac{\partial U}{\partial x}\,dt. \end{cases} \tag{5.22}$$

COROLLARY 5.3. *P is of* C*-type k if and only if it is of* B*-type k.*

[31] The non-vanishing conditions on the t_{ii} referred to in the preceding footnote become non-vanishing conditions on the h_{ii}.

For P is of C-type k if and only if the brace in (5.20) is a polynomial in x of degree $k + 1$. (5.20) thus reduces to (5.16), including the conditions (5.17).

There is no such close link between A-type and B-type. Consider, for example, the set

$$P_n(x) = \frac{x^n}{n!(n+1)!}.$$

It is of A-type *one* since $L[P_n] = P_{n-1}$, where

$$L[y] \equiv 2y' + xy''.$$

(Also, $J_0(t) = 2t$, $J_1(t) = t^2$, and $M_n(t) = (n+1)!t^n$.) On the other hand,

$$H(x, t) = \sum_0^\infty \frac{(xt)^n}{n!(n+1)!},$$

so that $\log H$ is decidedly not a polynomial in x. P is therefore of *infinite* B-type.

Let P be any set, and L its associated operator (i.e., $L[P_n] = P_{n-1}$). From (5.18) we have

$$U_1 L[P_n] + \cdots + U_n L^n[P_n] = nP_n,$$

so that we get

COROLLARY 5.4. *Every set P satisfies an equation of form*

$$V[y] \equiv \sum_{k=1}^{\infty} U_k L^k[y] = \lambda y, \tag{5.23}$$

where $\lambda = n$ for $y = P_n$. The U_n's are defined as in (5.18), *and L is the operator associated with P. If P is of* B-*type k, the coefficients in* (5.23) *are polynomials of maximum degree* $(k + 1)$.

In connection with sets of finite type (according to one definition or another) there arises the problem of the application of finite type sets to the solution of functional equations. This problem we reserve for another occasion. We shall terminate the present section by showing that the Legendre polynomials are of infinite type according to all the definitions given.

The B-type is determined by the maximum degree of the polynomials $T_n(x)$ of (5.11). Using relations (5.12) to (5.14), where $H = (1 - 2tx + t^2)^{-\frac{1}{2}}$, we find that

$$T(x, t) = \sum T_n(x)t^n = (1 - 2tx + t^2)^{-1}, \tag{a}$$

so that T_n satisfies the recurrence relation

$$T_n - 2xT_{n-1} + T_{n-2} = 0, \qquad n > 0. \tag{b}$$

With the initial conditions $T_0 = 1$, $T_1 = 2x$, the solution of (b) is

$$T_n(x) = \frac{1}{2\theta}\{(x+\theta)^{n+1} - (x-\theta)^{n+1}\}, \qquad \theta = (x^2 - 1)^{\frac{1}{2}}. \tag{c}$$

It is seen that $T_n(x)$ is of degree n, so that $\{X_n(x)\}$ is of infinite B-type (and C-type).

Now consider the A-type of $\{X_n\}$. If L is the associated operator then,

(d) $$L[X_n] = X_{n-1},$$

where

(e) $$L[y] \equiv L_0(x)y' + L_1(x)y'' + \cdots .$$

Multiplying (d) by t^n and summing from $n = 0$ to $n = \infty$, we obtain

(f) $$L[H] = tH, \qquad H = (1 - 2tx + t^2)^{-\frac{1}{2}}.$$

Relation (e), for $y = H$, simplifies to

(g) $$t = \sum_0^\infty L_n(x) \frac{1\cdot 3 \cdots (2n+1)t^{n+1}}{(1 - 2tx + t^2)^{n+1}};$$

and if we set

(h) $$\lambda = \frac{t}{1 - 2tx + t^2},$$

this becomes

(i) $$\frac{1}{2\lambda}\{1 + 2x\lambda - (1 + 4x\lambda + 4(x^2 - 1)\lambda^2)^{\frac{1}{2}}\} = \sum_0^\infty 1\cdot 3 \cdots (2n+1)\lambda^{n+1} L_n(x).$$

Since (g) is an identity in the variables t, x, so is (i) an identity in the variables λ, x. If $\{X_n\}$ is of finite A-type, the right member, and therefore the left, is a polynomial in x. This is manifestly untrue. Hence $\{X_n\}$ is of infinite A-type.

PENNSYLVANIA STATE COLLEGE.

ANNIHILATOR IDEALS AND REPRESENTATION ITERATION FOR ABSTRACT RINGS

By C. J. Everett, Jr.

1. **Introduction.** It has been shown[1] that in a ring Γ_0 with additive group M_0, every element g acting as a left multiplier on Γ_0, defines an operator λ in the endomorphism ring Ω_0 of M_0, such that $g\Gamma_0 = \lambda\Gamma_0$, $g\Gamma_0$ being the ordinary ring multiplication, and $\lambda\Gamma_0$ the map of M_0 by the operator λ. The set of all such operators λ forms a subring $\mathfrak{L} = \mathfrak{L}(\Gamma_0)$ of Ω_0 to which Γ_0 is ring homomorphic. We shall call $\mathfrak{L}$ the *left representation* of Γ_0. Similarly the right multipliers define a *right representation* $\mathfrak{R}^T$ of Γ_0, where $\mathfrak{R}^T$ consists of a ring inverse isomorphic to the subring $\mathfrak{R}$ of operators of Ω_0 defined by the right multipliers; $\mathfrak{R}^T$ being used in order that Γ_0 be ring homomorphic to $\mathfrak{R}^T$ in the ordinary sense. We have new rings $\mathfrak{L}$, $\mathfrak{R}^T$, each having an additive group with corresponding operator rings. We are thus free to iterate the process, forming left or right representations of representations in any order. We shall study these representation rings and their isomorphism to residue class rings of Γ_0 modulo certain annihilating ideals.[2]

2. **The ideal theory.** Let Γ be an abstract ring, and define Σ^m as the set of all products $s_1 \cdots s_m$ of m factors, $s_i \in \Sigma \subset \Gamma$. Denote by (r, l) the set of all x of Γ such that $\Gamma^r x \Gamma^l = 0$ $(r, l = 0, 1, \cdots)$. Γ^0 is merely deleted wherever it occurs formally. Thus $(0, 0) = 0$, and $(0, 1)\Gamma = 0$. It is clear that (r, l) is a 2-sided ideal in Γ and $(r, l) \subset (r + s, l + k)$ for all s, k. If Δ is a 2-sided ideal in Γ, we write $\Gamma - \Delta$ for the residue class ring of Γ mod Δ.

Let λ be the least l for which $(0, l) = (0, l + 1)$, ρ the least r such that $(r, 0) = (r + 1, 0)$. Γ is said to be of *l-type* λ, of *r-type* ρ. If $\lambda = 0$, Γ is called *l-definite*; if $\rho = 0$, *r-definite*; and if both, *definite*. Conditions on Γ for finiteness of type are given in §3. We shall consider only rings with finite ρ, λ.

Theorem 1. *If in* $\Gamma - (r, l)$, $(s, k) = 0$, *then in* $\Gamma - (r - i, l - j)$, $(i, j) = (i + s, j + k)$, *and conversely.*

Suppose $\Gamma^{i+s} x \Gamma^{j+k} \equiv 0$ in $\Gamma - (r - i, l - j)$; i.e., $\Gamma^{r+s} x \Gamma^{l+k} = 0$ in Γ and $\Gamma^s x \Gamma^k \equiv 0$ in $\Gamma - (r, l)$. Hence $x \equiv 0$ in the latter ring, and $\Gamma^r x \Gamma^l = 0$ in Γ. Hence in $\Gamma - (r - i, l - j)$, $\Gamma^i x \Gamma^j \equiv 0$, and in this ring $(i + s, j + k) \subset (i, j)$. The argument is reversible.

Received February 24, 1939.

[1] *Rings as groups with operators*, Bull. Amer. Math. Soc., vol. 45(1939), pp. 274–279.

[2] The extension of the theory to rings with rings of operators is readily made. Thus in the case of linear associative algebras $\mathfrak{L}$ and $\mathfrak{R}^T$ are the usual left and right regular matrix representations.

COROLLARY 1. *$\Gamma-(0, l)$ is l-definite if and only if $(0, l) = (0, l + 1)$. $\Gamma-(r, 0)$ is r-definite if and only if $(r, 0) = (r + 1, 0)$.*

COROLLARY 2. *$\Gamma-(0, \lambda)$ is l-definite, $\Gamma-(\rho, 0)$ is r-definite, and ρ, λ are the least such integers.*

THEOREM 2. *$(r, l) = (r, l + 1)$ implies $(r + s, l) = (r + s, l + 1)$ and $(r, l) = (r + 1, l)$ implies $(r, l + k) = (r + 1, l + k)$.*

This is immediate from definition of (r, l).

COROLLARY 3. *$\Gamma - (r, l)$ is definite if and only if $(r, l) = (r + 1, l + 1)$.*

The definiteness with Theorem 1 yields $(r, l) = (r, l + 1) = (r + 1, l)$, whence $(r, l) = (r + 1, l + 1)$ by Theorem 2. The converse follows from Theorem 1 after noticing that $(r, l) = (r + 1, l + 1)$ implies $(r, l) = (r, l + 1) = (r + 1, l)$.

COROLLARY 4. *$\Gamma-(\rho, \lambda)$ is definite.*

For $(0, \lambda) = (0, \lambda + 1)$ and $(\rho, 0) = (\rho + 1, 0)$ imply $(\rho, \lambda) = (\rho, \lambda + 1) = (\rho + 1, \lambda + 1)$, by Theorem 2.

THEOREM 3. *If Δ is a 2-sided ideal in Γ, then $\Gamma - \Delta$ l-definite implies $\Delta \supset (0, l)$; $\Gamma - \Delta$ r-definite implies $\Delta \supset (r, 0)$; and $\Gamma - \Delta$ definite implies $\Delta \supset (r, l)$, all r, l.*

Suppose in $\Gamma - \Delta$, $(0, 1) = (1, 0) = 0$; i.e., $x\Gamma \equiv 0\ (\Delta)$ implies $x \equiv 0\ (\Delta)$, and $\Gamma x \equiv 0\ (\Delta)$ implies $x \equiv 0\ (\Delta)$. Then for every r, l, $\Gamma^r x \Gamma^l \equiv 0\ (\Delta)$ implies $x \equiv 0\ (\Delta)$. Since $\Gamma^r(r, l)\Gamma^l = 0 \in \Delta$, $(r, l) \subset \Delta$. Similar proofs establish the rest of the theorem.

COROLLARY 5. *$(0, \lambda) = (0, \lambda + l)$, $(\rho, 0) = (\rho + r, 0)$, and $(s, k) = (s + r, k + l)$ for all s, k for which $\Gamma-(s, k)$ is definite.*

Use $\Delta = (0, \lambda)$, $(\rho, 0)$, (s, k) consecutively in Theorem 3.

COROLLARY 6. *$(0, \lambda)$, $(\rho, 0)$, (ρ, λ) are "minimum" ideals for which the corresponding residue class rings are l-definite, r-definite, and definite, respectively; i.e., each of these ideals is contained in every ideal having the corresponding property.*

We may now make several remarks concerning the lattice of ideals (r, l) in matric array, r, l serving as row and column indices, respectively. If $\Gamma-(r, l)$ is definite, (r, l) defines a *definite point* of the lattice. A definite point for which there exists no definite point (s, k), $s < r$, $k < l$, is a *frontier point*, the set of all such constituting the *frontier* of the lattice. The concept is significant in the iterated representation theory. Roughly, the frontier is the boundary between definite and non-definite points. It is to be noted that the first repetition in any row cannot occur to the right of the first in the preceding row. A similar remark applies to columns. All definite points define equal ideals.

Since in a commutative ring, $\lambda = \rho$, $(a, 0) = (0, a)$ and $(a + b, 0) = (a, b)$, it is seen that the frontier here is of step form, consisting of points (a, b) with $a + b = \lambda$, and $(c, 0)$, $(0, d)$, $c, d \geqq \lambda$. A nilpotent ring of index α ($\Gamma^\alpha = 0$) has a similar frontier with $a + b = \alpha - 1$, $c, d \geqq \alpha - 1$.

3. Iterated representation theory. Consider the sequence of homomorphic representation rings

$$(1) \qquad \Gamma_0 \sim \Gamma_1 \sim \cdots \sim \Gamma_n \sim \Gamma_{n+1} \sim \cdots \sim \Gamma_{n+p} \sim \cdots$$

obtained by taking successively left representations of left representations of Γ_0 as indicated in §1. We shall need a notation $L(\Sigma_1 : \Sigma_2)$ for the set of all elements g of a ring Γ such that $g\Sigma_1 \subset \Sigma_2$, where Σ_1, Σ_2 are subsets of Γ. If Σ_1 has the property that $\Gamma\Sigma_1 \subset \Sigma_1$, and if Σ_2 is a left ideal in Γ, $L(\Sigma_1 : \Sigma_2)$ is a 2-sided ideal in Γ. Thus $L(\Gamma^n : 0) = (0, n)$ in the previous notation, and $L(\Gamma_n : 0)$ is the set of all elements of Γ_n corresponding to 0 of Γ_{n+1} in (1); i.e., a ring Γ_n is l-definite in case $L(\Gamma_n : 0) = 0$, whence $\Gamma_n \cong \Gamma_{n+1}$. For convenience we define $L(\Gamma_n^0 : 0) = 0$. Note that for Σ_1, Σ_2 as above, $L(\Sigma_1^m : \Sigma_2)$ is a 2-sided ideal in Γ and $L(\Sigma_1^{m+n} : \Sigma_2) = L(\Sigma_1^m : L(\Sigma_1^n : \Sigma_2))$ since each set is contained in the other. We prove the

LEMMA. *In the correspondence* $\Gamma_n \sim \Gamma_{n+1}$ *of* (1), *the set in* Γ_n *corresponding to* $L(\Gamma_{n+1}^p : 0)$ *in* Γ_{n+1} *is* $L(\Gamma_n^{p+1} : 0)$.

For $p = 0$ the lemma is trivial under the convention $L(\Gamma_{n+1}^0 : 0) = 0$. Let $p = 1$. If $\xi \to \xi' \in L(\Gamma_{n+1} : 0)$, then $\xi\Gamma_n \to \xi'\Gamma_{n+1} = 0$. Hence $\xi \in L(\Gamma_n : L(\Gamma_n : 0)) = L(\Gamma_n^2 : 0)$. Conversely, if $\xi \in L(\Gamma_n^2 : 0)$ and $\xi \to \xi'$, $\xi\Gamma_n \to \xi'\Gamma_{n+1}$. But $\xi\Gamma_n \subset L(\Gamma_n : 0) \to 0$, and $\xi'\Gamma_{n+1} = 0$, whence $\xi' \in L(\Gamma_{n+1} : 0)$. Assuming the lemma for $p = k$, we prove it for $p = k + 1$. If $\xi \to \xi' \in L(\Gamma_{n+1}^{p+1} : 0) = L(\Gamma_{n+1} : L(\Gamma_{n+1}^p : 0))$, then $\xi\Gamma_n \to \xi'\Gamma_{n+1} \subset L(\Gamma_{n+1}^p : 0)$. Hence by hypothesis $\xi\Gamma_n \subset L(\Gamma_n^{p+1} : 0)$ and $\xi \in L(\Gamma_n : L(\Gamma_n^{p+1} : 0)) = L(\Gamma_n^{p+2} : 0)$. Conversely, if $\xi \in L(\Gamma_n^{p+2} : 0)$ and $\xi \to \xi'$, then $\xi\Gamma_n \to \xi'\Gamma_{n+1}$, where $\xi\Gamma_n \subset L(\Gamma_n^{p+1} : 0) \to L(\Gamma_{n+1}^p : 0)$. Hence we see that $\xi' \in L(\Gamma_{n+1} : L(\Gamma_{n+1}^p : 0)) = L(\Gamma_{n+1}^{p+1} : 0)$.

We have immediately

THEOREM 4. *In the homomorphism* $\Gamma_r \sim \Gamma_{r+n}$ *induced by* (1), $L(\Gamma_r^{p+n} : 0) \to L(\Gamma_{r+n}^p : 0)$ *and* $\Gamma_r - L(\Gamma_r^n : 0) \cong \Gamma_{r+n}$. *In particular*, $\Gamma_0 - L(\Gamma_0^n : 0) \cong \Gamma_n$.

COROLLARY 7. *If* $(0, l) = (0, l + 1)$ *in* Γ_0, *i.e.*, $\Gamma - (0, l)$ *l-definite, then* $(0, j) = (0, j + 1)$ *in* $\Gamma_0 - (0, l - j) \cong \Gamma_{l-j}$ *for all* j; *and conversely, if* $(0, j) = (0, j + 1)$ *in* $\Gamma - (0, l - j)$ *for any* j, $\Gamma - (0, l)$ *is l-definite.*

Let $i = s = r = 0$, $k = 1$ in Theorem 1.

THEOREM 5. *A necessary and sufficient condition that* $\Gamma_0 - (0, n)$ *be l-definite is that for some* l, k

$$L(\Gamma_k^l : 0) = L(\Gamma_k^{l+1} : 0);$$

i.e., in $\Gamma_0 - (0, k) \cong \Gamma_k$, $(0, l) = (0, l + 1)$.

THEOREM 6. *A sufficient condition that* $\Gamma_0 - (0, n)$ *be l-definite is that for some* k, $\Gamma_0 - (0, k) \cong \Gamma_k$ *have the maximal chain condition for 2-sided ideals.*

This is a weaker condition than the possession of the chain condition by Γ_0 itself. If $\Gamma_0 \sim \Gamma_1$ are any homomorphic rings, a chain condition on Γ_0 implies one on Γ_1, but not conversely, even for the rings here concerned. For example,

consider the null algebra Γ_0 ($\Gamma_0^2 = 0$) with infinite basis $(u_1, u_2, \cdots)$. Here $\Gamma_1 = 0$ has a chain condition, but not Γ_0, e.g., $(u_1) \subset (u_1 u_2) \subset \cdots$ is non-repeating.[3]

Rings with maximal chain condition for ideals are thus of finite l-type λ. In particular, if Γ is a linear associative algebra, the theory applies with obvious modifications, λ being $\leqq \alpha$, where α is the index of the algebra. As noted in §2, for a nilpotent algebra, $\lambda = \rho = \alpha - 1$, and $(\rho, 0) = (0, \lambda) = \Gamma$.

A strictly analogous theory may be deduced for right representations of right representations of Γ_0.[4] We shall now study the case of mixed iterates, left and right representations being arbitrarily derived from preceding ones.

Denote by $\mathfrak{L}(\Gamma)$ and $\mathfrak{R}^T(\Gamma)$ the left and right representations of a ring Γ. We may consider the iterated homomorphisms:

$$(2) \qquad \Gamma \sim \mathfrak{L}(\Gamma) \sim \mathfrak{R}^T(\mathfrak{L}(\Gamma)); \qquad \Gamma \sim \mathfrak{R}^T(\Gamma) \sim \mathfrak{L}(\mathfrak{R}^T(\Gamma)).$$

While $\mathfrak{R}^T(\mathfrak{L}(\Gamma))$ and $\mathfrak{L}(\mathfrak{R}^T(\Gamma))$ are conceptually distinct, $\mathfrak{R}(\mathfrak{L}(\Gamma))$, $\mathfrak{L}(\mathfrak{R}^T(\Gamma))$ being subrings of the endomorphism rings of the additive groups of $\mathfrak{L}(\Gamma)$ and $\mathfrak{R}^T(\Gamma)$, respectively, they have the property given by

THEOREM 7. $\mathfrak{L}(\mathfrak{R}^T(\Gamma)) \cong \mathfrak{R}^T(\mathfrak{L}(\Gamma)) \cong \Gamma - (1, 1)$.

This is a special case of the more general

THEOREM 8. *The ring resulting from the iteration of l left and r right representations in any mixed order is independent, up to isomorphism, of the order in which the iteration is performed, the resulting ring being always isomorphic with* $\Gamma - (r, l)$.

For in $\Gamma \sim \mathfrak{L}_1 \sim \mathfrak{L}_2 \sim \cdots \sim \mathfrak{L}_l \sim \mathfrak{R}^T(\mathfrak{L}_l) \sim \cdots \sim \mathfrak{R}_r^T(\mathfrak{L}_l)$, where $\mathfrak{L}_i$ denotes $\mathfrak{L}(\mathfrak{L}_{i-1})$ and $\mathfrak{R}_i^T(\mathfrak{L}_l)$ denotes $\mathfrak{R}^T(\mathfrak{R}_{i-1}^T(\mathfrak{L}_l))$, the set in $\mathfrak{L}_l$ corresponding to 0 in $\mathfrak{R}_r^T(\mathfrak{L}_l)$ is $R(\mathfrak{L}_l^r:0)$. Now in Γ, let $\xi \to \xi' \in R(\mathfrak{L}_l^r:0)$ in $\mathfrak{L}_l$. Then $\Gamma^r \xi \to \mathfrak{L}_l^r \xi' = 0$. Hence $\xi \in R(\Gamma^r : L(\Gamma^l:0)) = (r, l)$. Conversely, we may show that every element of $(r, l) \to 0$ in $\mathfrak{R}_r^T(\mathfrak{L}_l)$. The same result may be obtained for $\Gamma \sim \mathfrak{L}_l(\mathfrak{R}_r^T)$. Hence $r = l = 1$ gives Theorem 7, which in turn establishes the independence of order for Theorem 8. The significance of the frontier concept is now clear, it being necessary to take as many left and right representations (in any order) as will yield an ideal on or past the frontier, in order to obtain a definite representation.

4. Rings with units.

Such rings give rise to special theorems:

THEOREM 9. *If e_l is a left unit for Γ, then $(1, 0) = 0$ in Γ, $e_l + (0, 1)$ is the set of all left units of Γ, and e_l is unique if and only if Γ is l-definite.* $\Gamma - (0, \lambda)$ *is definite and has a principal unit.*

For $\Gamma x = 0$ implies $x = 0$ since $e_l \in \Gamma$. $(e_l + (0, 1))\, \Gamma \equiv \Gamma$ so every element of $e_l + (0, 1)$ is a left unit of Γ. Moreover, if e_l' is also a left unit of Γ, then $(e_l' - e_l)\Gamma \equiv \Gamma - \Gamma \equiv 0$ and $e_l' \equiv e_l$ modulo $(0, 1)$. Since for Γ, $\rho = 0$, $\Gamma - (0, \lambda)$ is

[3] I am indebted to George Whaples for this suggestion.

[4] In this case we use the notation $R(\Sigma_1:\Sigma_2)$ for all $g \in \Gamma$ such that $\Sigma_1 g \subset \Sigma_2$.

definite (Corollary 4). Hence the class $[e_l]$ is a unique left unit for this ring. It is thus a principal unit.[5]

THEOREM 10. *If* Γ *is r-definite, the frontier of* Γ *consists of* $(0, \lambda + i)$, *and* (r, λ), *all* i, r. *The frontier of a definite ring consists of* $(0, l)$, $(r, 0)$, *all* l, r.

Rings with left unit, and rings with unit afford examples of these two types.

5. **Representation of the center.** We shall now consider the relation between the center Z of Γ and the intersection $\mathfrak{L} \cap \mathfrak{R}$ of $\mathfrak{L}$ and $\mathfrak{R}$, these being subrings of the common over-ring Ω, the endomorphism ring of the additive group of Γ, as indicated in §1. Recall that $\Gamma \backsim \mathfrak{R}$ (inverse homomorphic: $r_1 \rightarrow \rho_1$, $r_2 \rightarrow \rho_2$ implies $r_1r_2 \rightarrow \rho_2\rho_1$).

THEOREM 11. $\mathfrak{L} \cap \mathfrak{R}$ *is in the center of* $\mathfrak{L}$ *and of* $\mathfrak{R}$ *and contains the map of* Z *under both* $\Gamma \sim \mathfrak{L}$ *and* $\Gamma \backsim \mathfrak{R}$.

Since $l(\Gamma r) = (l\Gamma)r$, it is clear that $\mathfrak{L}\mathfrak{R} = \mathfrak{R}\mathfrak{L}$ element-wise. Hence every element of $\mathfrak{L} \cap \mathfrak{R}$ is commutative with all the elements of $\mathfrak{L}$ and of $\mathfrak{R}$. Moreover, if z is in Z, $z\Gamma = \Gamma z$ and z defines the same operator whether acting as left or right multiplier, i.e., z maps into the same operator in both correspondences of the theorem, and its map is thus in $\mathfrak{L} \cap \mathfrak{R}$.

THEOREM 12. *If* Γ *is definite, the correspondences of Theorem* 11 *are one-one,* $\mathfrak{L} \cap \mathfrak{R}$ *is the center of* $\mathfrak{L}$ *and of* $\mathfrak{R}$, *and is the (left, right) representation of* Z.

Let $z \leftrightarrow \xi$ in the center of $\mathfrak{L}$. Because of the isomorphism, $\xi\mathfrak{L} \equiv \mathfrak{L}\xi$ implies $z\Gamma \equiv \Gamma z$, and z is in Z. By Theorem 11, $\xi \in \mathfrak{R}$. Hence the center of $\mathfrak{L}$ is $\mathfrak{L} \cap \mathfrak{R}$. Similarly for $\mathfrak{R}$.[6]

COROLLARY 8. *In a definite ring* Γ, $x\Gamma = \Gamma y$ *elementwise implies* $x = y$.

$x\Gamma = \Gamma y$ means that x and y correspond to the same operator θ which is thus in $\mathfrak{L}$ and in $\mathfrak{R}$, thus in $\mathfrak{L} \cap \mathfrak{R}$. x is therefore in Z, $x\Gamma = \Gamma x$, $\Gamma x = \Gamma y$, $\Gamma(y - x) = 0$, and $y = x$ since Γ is definite.

An *elementary* modul[7] may be defined, for the purposes of this note, as one for which there exists a ring Γ with unit element having the modul as additive group, and such that every endomorphism of the additive group is defined by a left multiplier.

COROLLARY 9. *The endomorphism ring of an elementary modul is commutative.*

Let $x \in \Gamma$; since x as right multiplier on Γ defines an endomorphism of the additive group, there exists a y such that $y\Gamma = \Gamma x$. By Corollary 8, x is in the center of Γ. But x was arbitrary.

UNIVERSITY OF WISCONSIN.

[5] If e is a unique left unit of a ring R, e is a right unit. For $(e + re - r)R \equiv R + rR - rR \equiv R$ implies $e + re - r = e$ for all $r \in R$. This proof is due to W. L. Mitchell.

[6] This is in agreement with the well-known fact that the left regular representation of a commutative linear associative algebra with unit element is identical with the transpose of the right regular representation.

[7] *Rings as groups with operators.*

PSEUDO-NORMED LINEAR SPACES AND ABELIAN GROUPS

By D. H. Hyers

In his study of linear topological spaces[1] J. von Neumann has introduced the notion of a pseudo-metric, which can be defined in any locally convex[2] linear topological space. It is a real-valued function and has many of the properties of a norm, including the fulfillment of the triangular inequality. The peculiarity of this pseudo-metric is that its value depends not only on the element x of the space but also on a neighborhood U out of a system of neighborhoods of the origin. In the present paper the pseudo-metric is generalized in the following way. First, the triangular inequality is replaced by a much weaker condition of continuity, and secondly, the neighborhood system is replaced by any "strongly" partially ordered set, and this makes it possible to postulate a "pseudo-norm" for a linear space without first postulating a topology.

In the first section of the paper pseudo-normed linear spaces are defined and shown to be the same as linear topological spaces. The pseudo-norm includes as special cases both the pseudo-metric of von Neumann and the pseudo-norm previously defined by the author in [3].

Pseudo-normed Abelian groups are defined in the second section of the paper, and it is shown that a topological Abelian group G has a "linear topological extension" if and only if the topology of G can be generated by a pseudo-norm. Thus pseudo-normed Abelian groups are just those Abelian groups which are topological subgroups of a linear topological space.

1. Let D be a set with elements $a, b, c, \cdots$ which has the composition property of Moore and Smith. That is, we are postulating that (i) if $a > b$, then $b \not> a$; (ii) if $a > b$ and $b > c$, then $a > c$; (iii) given a and b there exists c such that $c \geqq a$ and $c \geqq b$. The set D together with the relation $>$ will be called a ***strongly partially ordered space.***[3]

A linear space L will be said to be pseudo-normed with respect to D if there exists a real-valued function $n(x, d)$ defined on LD which satisfies the following postulates:

(1) $n(x, d) \geqq 0$; $n(x, d) = 0$ for all $d \in D$ implies $x = \theta$, where θ represents the zero of L;

(2) $n(\alpha x, d) = |\alpha| n(x, d)$ for all real α;

Received March 5, 1939; presented to the American Mathematical Society, September 6, 1938.

[1] See [5]. The numbers in brackets refer to the bibliography at the end of the paper.

[2] Local convexity is equivalent to von Neumann's convexity. See p. 158 of [6].

[3] Such sets were first used in general topology by G. Birkhoff. See [2].

(3) given $\eta > 0$, $e \in D$, there exist $\delta > 0$, $d \in D$ such that $n(x + y, e) < \eta$ for $n(x, d) < \delta$ and $n(y, d) < \delta$;

(4) $n(x, d) \geqq n(x, e)$ whenever $d > e$.

We shall call $n(x, d)$ the pseudo-norm of x with respect to d.

As an example of a pseudo-normed space which is not a normed space, take L to be the space of all real sequences $x = (x_1, x_2, x_3, \cdots)$ and D to be the set of positive integers. Define $n(x, d)$ by the equation $n(x, d) = \max_{1 \leq i \leq d} |x_i|$. It is clear that all the postulates for a pseudo-normed space are satisfied. Another example is the space[4] of all measurable functions $x(t)$ on the real open interval $(0, 1)$. If we take D to be the interval $(0, 1)$ with ordering as usual, we may define the pseudo-norm $n(x, d)$ by the equation

$$n(x, d) = \operatorname*{g.l.b.}_{m(E) \geq d} \{\operatorname*{l.u.b.}_{t \in E} |x(t)|\},$$

where E is a variable measurable subset of $(0, 1)$. In the above sequence space the pseudo-norm satisfies the triangular inequality, while in the space of measurable functions the pseudo-norm does not satisfy the triangular inequality, but only the weaker condition (3). In either case the topology given by the pseudo-norm, where x is a limit point of a set S if given $\delta > 0$, $d \in D$ there exists a point $y \in S$, $y \neq x$ such that $n(y - x, d) < \delta$, is equivalent to the usual metric topology for the space.[5]

By a linear topological space (l.t. space) we mean a linear Hausdorff space in which the fundamental operations $x + y$ and αx are continuous. It is known[6] that the following definition, essentially that of von Neumann, is equivalent to the above definition.

A linear space L will be called a l.t. space if there exists a fundamental system $\mathfrak{U}$ of subsets U of L such that the following postulates are satisfied.

I. The intersection of all the sets of $\mathfrak{U}$ is θ.

II. Given $U \in \mathfrak{U}$ and $V \in \mathfrak{U}$, there is a $W \in \mathfrak{U}$ such that $W \subset U \cdot V$.

III. Given $U \in \mathfrak{U}$, there is a $V \in \mathfrak{U}$ such that $\alpha V \subset U$ for all α satisfying $-1 \leqq \alpha \leqq 1$.

IV. For each $U \in \mathfrak{U}$ there is a $V \in \mathfrak{U}$ with $V + V \subset U$.

V. Given $x \in L$ and $U \in \mathfrak{U}$, there is a real number α such that $x \in \alpha U$.

These postulates are convenient in proving Theorem 1 below. The sets U, which are not necessarily open, will be called neighborhoods of θ. Any system $\mathfrak{U}'$ of sets will be called equivalent to $\mathfrak{U}$ if the Hausdorff equivalence criterion holds, that is, if for any $U \in \mathfrak{U}$ there is a $U' \in \mathfrak{U}'$ with $U' \subset U$ and conversely, given $V' \in \mathfrak{U}'$ there is a $V \in \mathfrak{U}$ with $V \subset V'$. Any system U' which is equivalent to $\mathfrak{U}$ clearly satisfies the Postulates I–V.

[4] Actually the elements of the space are equivalence classes of functions, where two functions are equivalent if they differ at most on a set of measure zero.

[5] See [1], pp. 9, 10, for the definitions of metrics in these two spaces.

[6] The equivalence is proved in [4].

Let L be a l.t. space which is pseudo-normed with respect to a strongly partially ordered space D. The pseudo-norm $n(x, d)$ will be said to *generate the topology of* L providing the sets[7] $U(d, \alpha) = [x; n(x, d) < \alpha]$, $d \in D$, $\alpha > 0$, form a system equivalent to the fundamental neighborhood system $\mathfrak{U}$.

THEOREM 1. *Every pseudo-normed linear space L is a linear topological space in which the pseudo-norm generates the topology of L. Conversely, given any linear topological space L, there exists a strongly partially ordered space D with respect to which L may be pseudo-normed in such a way that the pseudo-norm generates the topology of L.*

Proof. Let L be a pseudo-normed space and consider the subsets $U(d, \alpha) = [x; n(x, d) < \alpha]$. The intersection of all the $U(d, \alpha)$ is the origin by Postulates (1) and (2). Given $U(d, \alpha)$ and $U(e, \beta)$, let f be an element of D such that $f > d, f > e$, and let γ be a positive number less than α and β. For $x \in U(f, \gamma)$ we have $n(x, f) < \gamma$. From Postulate (4) it follows that $n(x, d) < \alpha$ and $n(x, e) < \beta$. Hence $U(f, \gamma) \subset U(d, \alpha) \cdot U(e, \beta)$, so that Postulate II is satisfied. The continuity of the sum (Postulate IV) follows from Postulate (3). It follows at once from Postulate (2) that Postulates III and V are satisfied.

Conversely, let L be a l.t. space, and let $\mathfrak{U}$ be the set of neighborhoods satisfying Postulates I–V. For each $U \in \mathfrak{U}$ and each $\alpha > 0$ let $V(\alpha, U)$ be the set of all elements of the form βx, where β ranges over the closed real interval $[-\alpha, \alpha]$ and x varies over U. Let $\mathfrak{U}'$ be the family of all the $V(\alpha, U)$. It follows from Postulate III and the continuity of the function βx that $\mathfrak{U}'$ is equivalent to $\mathfrak{U}$. The system $\mathfrak{U}'$ evidently has the following properties: (a) $V \in \mathfrak{U}'$ implies $\beta V \subset V$ for $|\beta| \leqq 1$; (b) if $V \in \mathfrak{U}'$ and $\beta \neq 0$, then $\beta V \in \mathfrak{U}'$; (c) $\mathfrak{U}'$ satisfies Postulates I–V. As the partially ordered space take the system $\mathfrak{U}'$ of neighborhoods, where $V_1 \geqq V_2$ means $V_1 \subset V_2$. Clearly $\mathfrak{U}'$ is strongly partially ordered by the relation $>$. For each $x \in L$ and each $V \in \mathfrak{U}'$ put $n(x, V) = \text{g.l.b. } \alpha > 0,\ x \in \alpha V$. Then $n(x, V)$ is obviously non-negative. Since $\mathfrak{U}'$ has property (a) it follows that $\beta > n(x, V)$ implies $x \in \beta V$. Hence if $n(x, V) = 0$ for all $V \in \mathfrak{U}'$, then $x \in V$ for all $V \in \mathfrak{U}'$ so that $x = \theta$ by Postulate I. This proves that (1) is satisfied. To show that (2) holds first take $\beta > 0$. Now

$$n(\beta x, V) = \text{g.l.b. } \alpha,\ \beta x \in \alpha V = \text{g.l.b. } \alpha,\ x \in \frac{\alpha}{\beta} V = \beta\left(\text{g.l.b. } \frac{\alpha}{\beta},\ x \in \frac{\alpha}{\beta} V\right) = \beta n(x, V).$$

For $\beta < 0$ a similar argument holds with a change of sign, since $V = -V$, and finally it is clear that $n(\theta, V) = 0$. In order to prove that Postulate (3) is satisfied let $\eta > 0$, $V_1 \in \mathfrak{U}'$ be given. Take $\eta_1 < \eta$. Since $\mathfrak{U}'$ satisfies Postulate IV and since $\eta_1 V_1 \in \mathfrak{U}'$ there exists $V_2 \in \mathfrak{U}'$ such that $V_2 + V_2 \subset \eta_1 V_1$. Now if $n(x, V_2) < 1$ and $n(y, V_2) < 1$ we have $x \in V_2$, $y \in V_2$ and therefore $n(x + y, V_1) \leqq \eta_1 < \eta$. To prove that $n(x, V)$ has the monotonic property (4) take $V_1 > V_2$, i.e., $V_1 \subset V_2$ where $V_1 \neq V_2$, $V_i \in \mathfrak{U}'$ $(i = 1, 2)$. For all $\alpha >$

[7] The symbol $[x;\quad]$ represents the set of x's having the property indicated after the semicolon.

$n(x, V_2)$ we have $x \in \alpha V_1 \subset \alpha V_2$. Hence $n(x, V_2) \geqq n(x, V_1)$. It remains to prove only that $\mathfrak{U}'$ is equivalent to the set of neighborhoods $U(V, \alpha) = [x; n(x, V) < \alpha]$, $\alpha > 0$. Given any set $U(V, \alpha)$ take $\beta < \alpha$. Then for any $x \in \beta V$ we have $n(x, V) \leqq \beta < \alpha$. Conversely, if $V \in \mathfrak{U}'$, then $n(x, V) < 1$ implies that $x \in V$. The proof of the theorem is now complete.

A pseudo-norm will be called triangular if it satisfies the triangular inequality

$$n(x + y, d) \leqq n(x, d) + n(y, d) \tag{3a}$$

for all x and y in L and all d in D. The inequality (3a) clearly implies Postulate (3). The pseudo-metric of von Neumann satisfies Postulates (1), (2), (3a) and (4), and hence is a triangular pseudo-norm.

Since l.t. spaces and pseudo-normed spaces are the same it is natural to try to express properties of l.t. spaces in terms of pseudo-norms. The next theorem affords a definition of local convexity in terms of pseudo-norms.

THEOREM 2. *A linear topological space is locally convex if and only if its topology can be generated by a triangular pseudo-norm.*

Proof. The necessity was proved by von Neumann on pp. 18, 19 of [5]. The sufficiency follows at once since the triangular inequality (3a) implies that the set $U(d, \alpha) = [x; n(x, d) < \alpha]$ is convex.

A pseudo-norm $n(x, d)$ which is independent of the element d of D reduces to the "pseudo-norm" defined in [3]. Hence in view of Theorem 3 of [3] we see that a l.t. space contains a bounded open set if and only if its topology is generated by a pseudo-norm independent of d. Finally it is obvious that a pseudo-norm is a norm in case it is both triangular and independent of d.

2. By a topological group we shall mean a group which is a Hausdorff space such that the group operation and the inverse operation are continuous. A topological Abelian group G (written additively) will be called a *topological subgroup* of a linear topological space L if G is a subgroup of L and if the original topology of G is equivalent to the topology for G obtained from L by relativization.[8] A l.t. space L which contains a topological subgroup G' which is continuously isomorphic[9] to a topological group G will be called a *linear topological extension of* G. Not every topological Abelian group containing no elements of finite period has a linear topological extension. For example, the additive group of all complex numbers topologized by means of the metric $\rho(x, y)$ where $\rho(x, y) = 1$ for $x \neq y$ and $\rho(x, x) = 0$ for all x is a topological group which has no l.t. extension.

Let G be an Abelian group, written additively, and let D be a strongly partially ordered space. The group G will be said to be *pseudo-normed* with respect

[8] I.e., the topology in which the open sets in G are the intersections of open sets in L with G.

[9] I.e., both the isomorphism $g \to g'$ and its inverse are continuous.

to D if there exists a real-valued function $n(g, d)$ defined on GD which satisfies the following postulates:

(A) $n(g, d) \geqq 0$; $n(g, d) = 0$ for all $d \in D$ implies $g = \theta$ where θ represents the zero of G;

(B) $n(\nu g, d) = |\nu| n(g, d)$ for any integer ν;

(C) given $\eta > 0$, $e \in D$, there exist $\delta > 0$, $d \in D$ such that $n(g, d) < \nu\delta$, $n(h, d) < \nu\delta$ imply $n(g + h, e) < \nu\eta$ for all positive integers ν.

(D) $n(g, d) \geqq n(g, e)$ whenever $d > e$.

From Postulates (A) and (B) it follows that $n(\theta, d) = 0$ for all $d \in D$ and that a pseudo-normed group has no elements of finite period. It is clear on the basis of Theorem 1 that any l.t. space is a pseudo-normed group. Every pseudo-normed group G is a topological group since the pseudo-norm generates a topology for G just as in the case of linear spaces.

THEOREM 3. *The topology of every topological subgroup of a linear topological space may be generated by a pseudo-norm. Conversely, every pseudo-normed Abelian group has a linear topological extension.*

Proof. The first statement follows at once from Theorem 1. To prove the converse we must construct a l.t. extension L for a given pseudo-normed group G.

We first extend the multiplier domain of G from the integers to the rationals. Consider the set of all triples (μ, ν, g), where μ and ν are integers, $\nu \neq 0$, and g is an element of G. Two triples (μ, ν, g) and (μ', ν', g') will be called equivalent (written $(\mu, \nu, g) \cong (\mu', \nu', g')$) in case $\mu\nu' g = \mu'\nu g'$. It is clear that this relation is reflexive, symmetric and transitive. We define the sum of two triples by the equation

$$(\mu, \nu, g) + (\rho, \tau, h) = (1, \nu\tau, \mu\tau g + \rho\nu h)$$

and the product of a rational number ρ/τ and a triple by the equation $(\rho/\tau)(\mu, \nu, g) = (\rho\mu, \tau\nu, g)$. It is easily shown that the equivalence relation $\cong$ is a congruence relation with respect to the operations of addition and of multiplication by rationals. That is, we can define the sum of two equivalence classes defined by the relation $\cong$ as the equivalence class of the sums, and the product of a rational and an equivalence class as the equivalence class of the products. The set of equivalence classes thus forms an Abelian group $\bar{G}$ with rational multipliers, and it is obvious that the subgroup consisting of those elements of $\bar{G}$ which have representatives of the form $(1, 1, g)$ is isomorphic with G. Since G is supposed to be pseudo-normed, we can define a pseudo-norm for the supergroup $\bar{G}$. Let x be any element of $\bar{G}$ and let (μ, ν, g) be any chosen representative of x. Let d be any element of the strongly partially ordered space D associated with G. Put $n(x, d) = |\mu/\nu| n(g, d)$. It is easily verified that $n(x, d)$ is independent of the representative of x selected. On making use of Postulates (A)–(D) on the pseudo-norm of G, one can show without difficulty that $n(x, d)$ satisfies Postulates (1)–(4) in §1 with the ex-

ception that in Postulate (2) the multiplier α is now restricted to be rational. Thus $\bar{G}$ is a pseudo-normed group which obviously contains a subgroup continuously isomorphic to G.

In order to obtain a l.t. extension of $\bar{G}$ we use the method of Cantor sequences. Define the relation between countable sequences out of $\bar{G}$ as follows: $(x_\mu) \sim (y_\mu)$ means that given $d \epsilon D$ and $\rho > 0$ there exists an integer μ_0 such that $\mu > \mu_0$ and $\nu > \mu_0$ imply $n(x_\mu - y_\nu, d) < \rho$. A sequence (x_μ) will be called fundamental if $(x_\mu) \sim (x_\mu)$. Clearly the relation $\sim$ is an equivalence relation among fundamental sequences. Let L denote the space whose elements are equivalence classes of fundamental sequences under the relation $\sim$. Define addition of two fundamental sequences by the equation $(x_\mu) + (y_\mu) = (x_\mu + y_\mu)$. Since the pseudo-norm for $\bar{G}$ satisfies Postulate (3) it follows that $(x_\mu + y_\mu)$ is a fundamental sequence and that the relation $\sim$ is a congruence relation with respect to addition. Hence we may define the sum $x + y$ of any two elements $x = \{(x_\mu)\}$ and $y = \{(y_\mu)\}$ of L as the equivalence class $\{(x_\mu + y_\mu)\}$. Let (α_μ) be a fundamental sequence of rational numbers and (x_μ) a fundamental sequence out of $\bar{G}$. Define $(\alpha_\mu) \circ (x_\mu) = (\alpha_\mu x_\mu)$. To demonstrate both that $(\alpha_\mu x_\mu)$ is a fundamental sequence and that $\sim$ is a congruence relation with respect to the operation $\circ$ we need only to prove that if (α_μ) is equivalent to (α'_μ) and $(x_\mu) \sim (x'_\mu)$, then $(\alpha_\mu x_\mu) \sim (\alpha'_\mu x'_\mu)$. Since $\bar{G}$ satisfies Postulate (3) it follows that for any chosen $e \epsilon D$ and $\eta > 0$ there is a $d \epsilon D$ and a $\rho > 0$ such that $n(y_i, d) < \rho$ $(i = 1, 2, 3)$ implies $n(y_1 + y_2 + y_3, e) < \eta$. Since (x'_μ) is fundamental there exists an integer ν such that $n(x'_\mu - x'_\nu, d) < \rho$ for $\mu > \nu$. Now let $\nu_1 > \nu$ be chosen so large that $\lambda > \nu_1$ and $\mu > \nu_1$ imply that

$$n(\alpha_\lambda(x_\lambda - x'_\mu), d) < \rho \quad \text{and} \quad |\alpha_\lambda - \alpha'_\mu| < \min\left[1, \frac{\rho}{n(x'_\nu, d)}\right].$$

Since $\alpha_\lambda x_\lambda - \alpha'_\mu x'_\mu = \alpha_\lambda(x_\lambda - x'_\mu) + (\alpha_\lambda - \alpha'_\mu)(x'_\mu - x'_\nu) + (\alpha_\lambda - \alpha'_\mu)x'_\nu$, we see that $n(\alpha_\lambda x_\lambda - \alpha'_\mu x'_\mu, e) < \eta$ for $\lambda > \nu_1$, $\mu > \nu_1$. A real number according to the Cantor construction is an equivalence class of fundamental sequences of rationals. Since the relation $\sim$ is a congruence relation with respect to the operation $\circ$, we may define the product αx of an element $x = \{(x_\lambda)\}$ of L by a real number $\alpha = \{(\alpha_\lambda)\}$ as the element $\{(\alpha_\lambda x_\lambda)\}$ of L. With these definitions of addition and of multiplication by reals it is clear that L is a linear space. Let $U(d, \rho)$ be the set of all points x of L for which there exists a representative (x_τ) of x and an integer σ such that $n(x_\tau, d) < \rho$ for all $\tau > \sigma$. We shall now show that the set $\mathfrak{U}$ of all the sets $U(d, \rho)$ satisfies Postulates I–V in §1, and this will prove that L is a l.t. space. That U satisfies I, II and IV follows without difficulty. To show that III is fulfilled let x be any point of $U(d, \rho)$ and let (x_τ) be a representative of x with $n(x_\tau, d) < \rho$ for $\tau > \sigma$. The pseudo-norm for $\bar{G}$ satisfies Postulate (2) for rational multipliers. Hence $n(-x_\tau, d) < \rho$ for $\tau > \sigma$ so that $-x = \{(-x_\tau)\} \epsilon U(d, \rho)$. For any real α in the interval $-1 < \alpha < 1$ let (α_τ) be a sequence of rationals converging to α. Then for some $\omega > \sigma$ it is true that $\tau > \omega$ implies that $|\alpha_\tau| < 1$. Hence for $\tau > \omega$ we have $n(\alpha_\tau x_\tau, d) < \rho$.

Therefore for all α satisfying $-1 \leqq \alpha \leqq 1$ we have $\alpha U(d, \rho) \subset U(d, \rho)$. To demonstrate that V is fulfilled let $x \in L$, $e \in D$ and $\eta > 0$ be given. Let $d \in D$ and $\delta > 0$ be chosen in accordance with Postulate (3) for pseudo-norms. Let (x_τ) be any representative of x, and choose σ so that $\tau > \sigma$ implies that $n(x_\tau - x_\sigma, d) < \delta$. Let β be a positive rational $< \min(1, \delta/n(x_\sigma, d))$. Since $\beta x_\tau = \beta(x_\tau - x_\sigma) + \beta x_\sigma$ and since $n(\beta(x_\tau - x_\sigma), d) < \delta$ and $n(\beta x_\sigma, d) < \delta$ for $\tau > \sigma$, it follows from the choice of d and δ that $n(\beta x_\tau, e) < \eta$ for all $\tau > \sigma$. That is, $\beta x \in U(e, \eta)$ where $\beta \neq 0$. On taking $\alpha = 1/\beta$ Postulate V is seen to be satisfied.

We have shown that L is a l.t. space. Moreover, the pseudo-normed group $\bar{G}$ is clearly continuously isomorphic to the subgroup of L consisting of those equivalence classes containing a representative (x_τ) such that $x_\tau = x_1$ for all τ. Since the original group G is continuously isomorphic to a subgroup of $\bar{G}$, it follows that L is a l.t. extension of G. This completes the proof of the theorem.

The extension L of G is not in general a complete space. If a complete extension is desired, the group $\bar{G}$ should be extended as described in [2] by G. Birkhoff, using directed sets in the place of countable sequences.

References

1. S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932.
2. G. Birkhoff, *Moore-Smith convergence in general topology*, Annals of Mathematics, vol. 38(1937), pp. 37–56.
3. D. H. Hyers, *A note on linear topological spaces*, Bulletin of the American Mathematical Society, vol. 44(1938), pp. 76–81.
4. D. H. Hyers, *Locally bounded linear topological spaces*, to appear in Revista des Ciencas, University of Lima.
5. J. von Neumann, *On complete topological spaces*, Transactions of the American Mathematical Society, vol. 37(1935), pp. 1–20.
6. J. V. Wehausen, *Transformations in linear topological spaces*, this Journal, vol. 4(1938), pp. 157–169.

The University of Wisconsin.

A MEAN ERGODIC THEOREM

By Nelson Dunford

In this note we shall give an ergodic theorem of the von Neumann type[1] for continuous n-parameter uniformly bounded semi-groups of linear transformations in a Banach space. The method of proof is an extension of one used by F. Riesz[2] and K. Yosida[3] for the discrete case in one dimension.

Throughout the note we shall use the following notation and terminology. E_n is Euclidean n-space, E_n^+ is the set of points $\alpha = (a_1, \cdots, a_n)$ in E_n with $a_i \geqq 0$ $(i = 1, \cdots, n)$, I_r is a cube of side $r > 0$ in E_n. That is, I_r is the set of points $\alpha = (a_1, \cdots, a_n) \in E_n$ for which

$$c_r^j \leqq a_j \leqq c_r^j + r \qquad (j = 1, \cdots, n), \tag{1}$$

where c_r^j $(j = 1, \cdots, n)$ is an arbitrary real function defined for $r > 0$. When we are dealing with functions defined only on E_n^+, it will be assumed that all cubes mentioned are in E_n^+. We shall sometimes use the symbol V_r for the volume r^n of a cube I_r. A set T_α $(\alpha \in E_n^+)$ of linear transformations in a Banach space $\mathfrak{B}$ is said to form a *semi-group* in case

$$T_{\alpha+\beta} = T_\alpha T_\beta, \qquad \alpha, \beta \in E_n^+, \tag{2}$$

and in case T_α is defined for every $\alpha \in E_n$, it is said to form a *group* if T_0 is the identity and equation (2) holds for all $\alpha, \beta \in E_n$. A group (or semi-group) T_α is said to be *uniformly bounded* in case $\| T_\alpha \|$ is bounded in α, and it is said to be *weakly measurable* in case the numerical function $\bar{x} T_\alpha x$ is measurable (in the sense of Lebesgue) for every $x \in \mathfrak{B}$ and $\bar{x} \in \bar{\mathfrak{B}}$ (the space conjugate to $\mathfrak{B}$). A function x_α defined on E_n (or E_n^+) is said to be *almost separably valued* in case there is a set E_0 of measure zero such that the set of points x_α, $\alpha \bar{\in} E_0$, is a separable subset of $\mathfrak{B}$. If for every cube I_r of side $r > 0$ we have a point $y(I_r) \in \mathfrak{B}$, then the set of all $y(I_r)$ is said to be *weakly compact with respect to* $r \to \infty$ in case every particular set I_r $(0 < r < \infty)$ of cubes contains a sequence I_{r_p} with $r_p \to \infty$ and $y(I_{r_p})$ converging weakly to an element of $\mathfrak{B}$. A function x_α defined on a cube is said to be *measurable* in case it is the limit almost every-

Received March 18, 1939. The results in this paper were communicated in the note *An ergodic theorem for n-parameter groups*, Proc. Nat. Acad. Sci., vol. 25(1939), pp. 195–196.

[1] J. von Neumann, *Proof of the quasi ergodic hypothesis*, Proc. Nat. Acad., vol. 18(1932), pp. 70–82.

[2] F. Riesz, *Some mean ergodic theorems*, Jour. London Math. Soc., vol. 13(1938), pp. 274–278.

[3] K. Yosida, *Mean ergodic theorem in Banach spaces*, Proc. Imp. Acad. Tokyo, vol. 14(1938), pp. 292–294.

where of step functions, and it is said to be *absolutely integrable*[4] in case it is measurable and $\| x_\alpha \|$ is integrable in the sense of Lebesgue. Finally, since most of what we shall say is true for semi-groups on E_n^+ as well as for groups on E_n, we shall use the letter G to stand for either E_n or E_n^+; so that a statement concerning a semi-group T_α on G is meant to cover both cases.

THEOREM 1. *Let $\mathfrak{B}$ be a Banach space and T_α ($\alpha \in G$) a weakly measurable and uniformly bounded semi-group of linear transformations in $\mathfrak{B}$. If for each x in $\mathfrak{B}$ the function $T_\alpha x$ is almost separably valued, then $T_\alpha x$ is absolutely integrable over every set $E \subset G$ of finite measure. Let $\mathfrak{M}$ be the set of those $x \in \mathfrak{B}$ for which*

$$\text{(i)} \qquad y_r = \frac{1}{V_r} \int_{I_r} T_\alpha x \, d\alpha$$

is weakly compact with respect to $r \to \infty$. Then $\mathfrak{M}$ is a closed linear manifold in $\mathfrak{B}$ and there is a continuous linear mapping $y = Ux$ of $\mathfrak{M}$ into itself with $U^2 = U$ and such that

$$\text{(ii)} \qquad T_\alpha U x = Ux, \qquad x \in \mathfrak{M},\ \alpha \in G,$$

$$\text{(iii)} \qquad Ux = \lim_{r \to \infty} \frac{1}{V_r} \int_{I_r} T_\alpha x \, d\alpha, \qquad x \in \mathfrak{M},$$

(iv) *$U\mathfrak{M}$ is the closed linear manifold consisting of those x in $\mathfrak{B}$ for which $T_\alpha x = x$, $\alpha \in G$.*

If the space $\mathfrak{B}$ satisfies any one of the following conditions:

(a) *its unit sphere is weakly compact,*

(b) *it is reflexive,*[5]

(c) *it has an equivalent uniformly convex norm;*[6]

then $\mathfrak{M} = \mathfrak{B}$. The limit in (iii) *fails to exist for $x \bar{\in} \mathfrak{M}$.*

It is a result of Gelfand[7] and Pettis[8] that the measurability of $T_\alpha x$ is a consequence of its weak measurability together with the fact that it is almost separably valued. Since $\| T_\alpha x \| \leqq C \| x \|$, where C is a bound for $\| T_\alpha \|$, the function $T_\alpha x$ is then absolutely integrable over every set $E \subset G$ of finite measure.

[4] For a discussion of the type of integration of vector valued functions that we use here the reader is referred to S. Bochner, Fund. Math., vol. 20(1933), pp. 262–276; and N. Dunford, Trans. Amer. Math. Soc., vol. 37(1935), pp. 441–453, corrections, ibid., vol. 38(1936), pp. 600–601.

[5] A space is called reflexive in case to each $\bar{\bar{x}} \in \bar{\bar{\mathfrak{B}}}$ there is an $x \in \mathfrak{B}$ such that $\bar{\bar{x}}\bar{x} = \bar{x}x$ for every $\bar{x} \in \bar{\mathfrak{B}}$. This terminology is due to E. R. Lorch.

[6] A space is said to be uniformly convex in case the length of a chord joining two points on the surface of the unit sphere approaches zero as its center point approaches the surface of the sphere. This concept is due to J. A. Clarkson.

[7] I. Gelfand, *Zur Theorie abstrakter Funktionen*, Comptes Rendus de l'Acad. des Sc. de l'URSS, vol. 17(1937), pp. 243–245.

[8] B. J. Pettis, *On integration in vector spaces*, Trans. Amer. Math. Soc., vol. 44(1938), pp. 277–304; in particular Theorem 1.1.

We shall leave the verification of the properties stated for $\mathfrak{M}$ until later, and fix our attention for the present on a single point x in $\mathfrak{M}$. We start with an arbitrary fixed set (1) of cubes I_r. There is a subsequence y_{r_p} of y_r such that $r_p \to \infty$ and $\bar{x}y_{r_p} \to \bar{x}y$ for some $y \in \mathfrak{B}$ and every $\bar{x} \in \bar{\mathfrak{B}}$. We shall express these facts by

$$y_{r_p} \underset{w}{\rightarrow} y. \tag{3}$$

From (3) and the continuity of T_α we have

$$T_\alpha y_{r_p} \underset{w}{\rightarrow} T_\alpha y, \qquad \alpha \in G. \tag{4}$$

We wish next to show that

$$T_\alpha y = y, \qquad \alpha \in G. \tag{5}$$

It will be sufficient in establishing (5) to show that it holds for every vector $\alpha^j = (0, 0, \cdots, a_j, \cdots, 0)$ with all components zero except the j-th; for if $\alpha = (a_1, \cdots, a_n)$, then $\alpha = \alpha^1 + \cdots + \alpha^n$ and $T_\alpha = T_{\alpha^1}T_{\alpha^2} \cdots T_{\alpha^n}$. Now[9]

$$T_{\alpha^j} y_{r_p} = \frac{1}{r_p^n} \int_{I_{r_p}} T_{\beta+\alpha^j} x \, d\beta = \frac{1}{r_p^n} \int_{I_{r_p}+\alpha^j} T_\beta x \, d\beta,$$

so that

$$\begin{aligned} T_{\alpha^j} y_{r_p} - y_{r_p} &= \frac{1}{r_p^n} \left[\int_{I_{r_p}+\alpha^j} T_\beta x \, d\beta - \int_{I_{r_p}} T_\beta x \, d\beta \right] \\ &= \frac{1}{r_p^n} \left[\int_A T_\beta x \, d\beta - \int_B T_\beta x \, d\beta \right], \end{aligned}$$

where

$$\begin{aligned} A &= (I_{r_p} + \alpha^j) - I_{r_p}(I_{r_p} + \alpha^j), \\ B &= I_{r_p} - I_{r_p}(I_{r_p} + \alpha^j). \end{aligned}$$

Hence $|B| = |A| = r_p^{n-1} |a_j|$ if $r_p > |a_j|$. Since $\| T_\beta x \| \leqq C \| x \|$, we have then

$$\| T_{\alpha^j} y_{r_p} - y_{r_p} \| \leqq 2 \cdot C \cdot \| x \| \cdot | a_j | \cdot r_p^{-1} \qquad \text{if } r_p > | a_j |.$$

By (3) and (4), $T_{\alpha^j} y_{r_p} - y_{r_p} \underset{w}{\rightarrow} T_{\alpha^j} y - y$, and so we must have

$$\| T_{\alpha^j} y - y \| \leqq \liminf_{p \to \infty} \| T_{\alpha^j} y_{r_p} - y_{r_p} \| = 0,$$

and thus equation (5) is established.

[9] Here we are using the fact that $T_\alpha \int_I T_\beta x \, d\beta = \int_I T_{\alpha+\beta} x \, d\beta$. See Theorem 2.3 in the authors paper *Integration and linear operations*, Trans. Amer. Math. Soc., vol. 40(1936), pp. 474–494.

Now let $a > 0$ be a fixed positive number and $\mathfrak{M}_a^j$ the closed linear manifold in $\mathfrak{B}$ determined by the points

$$T_{\alpha+a^j}x - T_\alpha x, \qquad \alpha \epsilon G,$$

where α^j is the vector with a in the j-th place and all other components zero. Let $\mathfrak{M}_a = [\mathfrak{M}_a^1, \cdots, \mathfrak{M}_a^n]$, i.e., $\mathfrak{M}_a$ is the closed linear manifold in $\mathfrak{B}$ determined by the manifolds $\mathfrak{M}_a^j$ $(j = 1, \cdots, n)$. It will next be shown that the vector

$$y - V_a^{-1} \int_{J_a} T_\beta x \, d\beta \tag{6}$$

belongs to $\mathfrak{M}_a$ for any cube J_a of side a. To show this it will suffice to prove that $\bar{x}\left[y - V_a^{-1} \int_{J_a} T_\beta x \, d\beta\right] = 0$ for every vector $\bar{x} \epsilon \bar{\mathfrak{B}}$ which vanishes on $\mathfrak{M}_a$. Accordingly let $\bar{x}$ be such a functional, then

$$\bar{x}T_{\alpha+a^j}x = \bar{x}T_\alpha x \qquad (\alpha \epsilon G; j = 1, \cdots, n). \tag{7}$$

Let $\left[\frac{r}{a}\right]$ be the greatest integer in $\frac{r}{a}$ and let J_a be any cube of side a. In view of (7) we have

$$\frac{1}{V_r} \int_{I_r} \bar{x}T_\beta x \, d\beta = \frac{1}{V_r}\left\{\left[\frac{r}{a}\right]^n \int_{J_a} \bar{x}T_\beta x \, d\beta + \int_{R_r^a} \bar{x}T_\beta x \, d\beta\right\}, \tag{8}$$

where $R_r^a = \sum_{j=1}^n R_{r,j}^a$ and $R_{r,j}^a$ is the rectangle defined as

$$c_r^i \leqq a_i \leqq c_r^i + r \qquad (i = 1, \cdots, n; i \neq j),$$

$$a\left[\frac{r}{a}\right] + c_r^j \leqq a_j \leqq c_r^j + r.$$

Since $r - \left[\frac{r}{a}\right]a < a$, we have $| R_r^a | \leqq \sum_{j=1}^n | R_{r,j}^a | \leqq nar^{n-1}$, and thus

$$\left\| \frac{1}{V_r} \int_{R_r^a} \bar{x}T_\beta x \, d\beta \right\| \leqq n \cdot a \cdot C \cdot \| \bar{x} \| \cdot \| x \| \cdot r^{-1}. \tag{9}$$

Upon combining (i), (3), (8) and (9) with the fact that

$$\frac{1}{V_r}\left[\frac{r}{a}\right]^n = \frac{1}{a^n}\left[\frac{r}{a}\right]^n \left(\frac{a}{r}\right)^n \to \frac{1}{V_a} \quad \text{as} \quad r \to \infty,$$

we have $\bar{x}y = V_a^{-1} \int_{J_a} \bar{x}T_\beta x \, d\beta$, i.e.,

$$\bar{x}\left[y - V_a^{-1} \int_{J_a} T_\beta x \, d\beta\right] = 0.$$

This shows that the vector (6) belongs to $\mathfrak{M}_a$. For a given $\epsilon > 0$ we can find therefore a point $z \in \mathfrak{M}_a$, $\| z \| < \epsilon$, a finite number of constants b_{ij}, and a finite number of vectors $\alpha_{ij} \in G$ such that

$$y - \frac{1}{V_a} \int_{J_a} T_\beta x \, d\beta = \sum_{j=1}^{q} \sum_{i=1}^{p} b_{ij}[T_{\alpha_{ij}+\alpha^j} x - T_{\alpha_{ij}} x] + z. \tag{10}$$

Upon operating on both sides of (10) by T_γ and averaging over I_r, we obtain (see (5))

$$\begin{aligned} y - \frac{1}{V_r} \int_{I_r} T_\gamma \left(\frac{1}{V_a} \int_{J_a} T_\beta x \, d\beta \right) d\gamma &= \frac{1}{V_r} \int_{I_r} T_\gamma z \, d\gamma \\ &+ \sum_{j=1}^{q} \sum_{i=1}^{p} b_{ij} \frac{1}{V_r} \left[\int_{I_r} T_{\gamma+\alpha_{ij}+\alpha^j} x \, d\gamma - \int_{I_r} T_{\gamma+\alpha_{ij}} x \, d\gamma \right]. \end{aligned} \tag{11}$$

Now by a change of variable of integration

$$\int_{I_r} T_{\gamma+\alpha_{ij}+\alpha^j} x \, d\gamma = \int_{I_r+\alpha^j} T_{\gamma+\alpha_{ij}} x \, d\gamma,$$

and so

$$\int_{I_r} T_{\gamma+\alpha_{ij}+\alpha^j} x \, d\gamma - \int_{I_r} T_{\gamma+\alpha_{ij}} x \, d\gamma = \int_{A_1} T_{\gamma+\alpha_{ij}} x \, d\gamma - \int_{B_1} T_{\gamma+\alpha_{ij}} x \, d\gamma,$$

where

$$\begin{aligned} A_1 &= (I_r + \alpha^j) - I_r(I_r + \alpha^j), \\ B_1 &= I_r - I_r(I_r + \alpha^j). \end{aligned}$$

Thus for $r > a$ we have $| A_1 | = | B_1 | = ar^{n-1}$, and this shows that

$$\left\| y - \frac{1}{V_r} \int_{I_r} T_\gamma \left(\frac{1}{V_a} \int_{J_a} T_\beta x \, d\beta \right) d\gamma \right\| \leqq [2 \cdot B \cdot a \cdot \| x \| \cdot r^{-1} + \epsilon] C, \quad r > a$$

(if we make use of (11) and the fact that $\| z \| < \epsilon$), where $B = \sum_{j=1}^{q} \sum_{i=1}^{p} | b_{ij} |$. Hence

$$\limsup_{r \to \infty} \left\| y - \frac{1}{V_r} \int_{I_r} T_\gamma \left(\frac{1}{V_a} \int_{J_a} T_\beta x \, d\beta \right) d\gamma \right\| \leqq C \cdot \epsilon,$$

and since $\epsilon > 0$ was arbitrary, we have

$$- \lim_{r \to \infty} \frac{1}{V_r} \int_{I_r} T_\gamma u \, d\gamma, \tag{12}$$

for every vector u of the form

$$u = \frac{1}{V_a} \int_{J_a} T_\beta x \, d\beta, \tag{13}$$

where in (13) a is an arbitrary positive number and J_a is an arbitrary cube of side a.

Now let $\epsilon > 0$ be arbitrary and fix β_0 in the set where $\int_J T_\beta x \, d\beta$ is differentiable to its integrand.[10] Let J_a be a fixed cube of side $a > 0$ such that $\beta_0 \epsilon J_a$ and

$$\left\| \frac{1}{V_a} \int_{J_a} T_\beta x \, d\beta - T_{\beta_0} x \right\| < \epsilon. \tag{14}$$

Then

$$\begin{aligned} \left\| y - \frac{1}{V_r} \int_{I_r} T_{\beta+\beta_0} x \, d\beta \right\| &\leqq \left\| y - \frac{1}{V_r} \int_{I_r} T_\gamma \left(\frac{1}{V_a} \int_{J_a} T_\beta x \, d\beta \right) d\gamma \right\| \\ &\quad + \frac{1}{V_r} \left\| \int_{I_r} T_\gamma \left(\frac{1}{V_a} \int_{J_a} T_\beta x \, d\beta \right) d\gamma - \int_{I_r} T_{\gamma+\beta_0} x \, d\gamma \right\|, \end{aligned} \tag{15}$$

and by (14) we have

$$\begin{aligned} &\left\| \int_{I_r} T_\gamma \left(\frac{1}{V_a} \int_{J_a} T_\beta x \, d\beta \right) d\gamma - \int_{I_r} T_{\gamma+\beta_0} x \, d\gamma \right\| \\ &\qquad = \left\| \int_{I_r} T_\gamma \left\{ \frac{1}{V_a} \int_{J_a} T_\beta x \, d\beta - T_{\beta_0} x \right\} d\gamma \right\| \leqq V_r \cdot C \cdot \epsilon. \end{aligned}$$

Thus by (12) and (15)

$$\limsup_{r \to \infty} \left\| y - \frac{1}{V_r} \int_{I_r} T_{\beta+\beta_0} x \, d\beta \right\| \leqq C \cdot \epsilon.$$

Since $\epsilon > 0$ was arbitrary, we have

$$y = \lim_{r \to \infty} T_{\beta_0} \frac{1}{V_r} \int_{I_r} T_\beta x \, d\beta. \tag{16}$$

If T_α were a group instead of a semi-group, we could operate on both sides of (16) with the inverse $T_{-\beta_0}$ of T_{β_0} and obtain the desired result. As it is, however, a further argument is necessary. Let us suppose that $\beta_0 = (b_1, \cdots, b_n)$ has been fixed so that $0 \leqq b_i \leqq 1$ $(i = 1, \cdots, n)$. Let I'_r be those points of I_r whose distance to the boundary of G is $\geqq 1$. Then

$$| I_r - I'_r | \leqq r^n - (r-1)^n, \tag{17}$$

and $I'_r - \beta_0 \subset G$. From (17) we see that

$$\left\| \frac{1}{V_r} \int_{I_r} T_\beta x \, d\beta - \frac{1}{V_r} \int_{I'_r} T_\beta x \, d\beta \right\| \leqq [1 - (1 - r^{-1})^n] \cdot C \cdot \| x \|. \tag{18}$$

[10] This is true for almost all values of β_0. Actually one can prove easily from the group property on the T_α that this integral is differentiable to its integrand at all interior points of G.

Now

$$(19)\qquad \left\| y - \frac{1}{V_r}\int_{I_r'} T_\beta x\, d\beta \right\| \leqq \left\| y - T_{\beta_0}\frac{1}{V_r}\int_{I_r'} T_\beta x\, d\beta \right\| + \frac{1}{V_r}\left\| T_{\beta_0}\left\{\int_{I_r'} T_\beta x\, d\beta - \int_{I_r'-\beta_0} T_\beta x\, d\beta\right\} \right\|.$$

It follows from (16) and (18) that

$$(20)\qquad \lim_{r\to\infty}\left\| y - T_{\beta_0}\frac{1}{V_r}\int_{I_r'} T_\beta x\, d\beta \right\| = 0.$$

If we write $A = I_r' - I_r'(I_r' - \beta_0)$, $B = (I_r' - \beta_0) - I_r'(I_r' - \beta_0)$, then the measures $|A|$, $|B|$ of A, B are at most nr^{n-1} and

$$(21)\qquad \frac{1}{V_r}\left\| T_{\beta_0}\left\{\int_{I_r'} T_\beta x\, d\beta - \int_{I_r'-\beta_0} T_\beta x\, d\beta\right\} \right\| = \frac{1}{V_r}\left\| T_{\beta_0}\left\{\int_{A} T_\beta x\, d\beta - \int_{B} T_\beta x\, d\beta\right\} \right\| \leqq 2\cdot n\cdot C^2\cdot \|x\|\cdot r^{-1}.$$

Thus (21) and (20) show that the left side of (19) approaches zero, and this fact combined with (18) shows that

$$(22)\qquad y = \lim_{r\to\infty}\frac{1}{V_r}\int_{I_r} T_\beta x\, d\beta.$$

Upon placing $y = Ux$, $y_r = U_r x$, we see that U_r is a uniformly bounded set of continuous linear operators with $\lim_{r\to\infty} U_r x = Ux$ for x in $\mathfrak{M}$; thus $U_r x$ converges for every x in $\overline{\mathfrak{M}}$. This proves that $\mathfrak{M}$ is closed and that U is continuous on $\mathfrak{M}$. Since $T_\alpha Ux = Ux$ (see (5)), we have $U_r Ux = Ux$, which equality shows that $U\mathfrak{M} \subset \mathfrak{M}$ and that $U^2 = U$. Statement (iv) of the theorem follows immediately from the preceding remarks.

If the unit sphere in $\mathfrak{B}$ is weakly compact, then every bounded set is weakly compact and thus $\mathfrak{M} = \mathfrak{B}$. If $\mathfrak{B}$ is reflexive, then its unit sphere is weakly compact[11] and so in this case also $\mathfrak{M} = \mathfrak{B}$. If $\mathfrak{B}$ has an equivalent uniformly convex norm, then $\mathfrak{B}$ is reflexive,[12] and so in this case also $\mathfrak{M} = \mathfrak{B}$. This completes the proof of Theorem 1.

An instance of Theorem 1 will now be given. Let E be an abstract set of points P and suppose there is a finite completely additive non-negative measure

[11] For a proof of this fact see B. J. Pettis, *A note on regular Banach spaces*, Bull. Amer. Math. Soc., vol. 44(1938), pp. 420–428, and V. Gantmakner and V. Šmulian, *Sur les espaces dont la sphère unitaire est faiblement compacte*, Comptes Rendus de l'Acad. des Sc. de l'URSS, vol. 17(1937), pp. 91–94.

[12] For a proof of this see B. J. Pettis, *A proof that every uniformly convex space is reflexive*, this Journal, vol. 5(1939), pp. 249–253; and D. Milman, *On some criteria for the regularity of spaces of* (B), Comptes Rendus de l'Acad. des Sc. de l'URSS, vol. 20(1938), pp. 243–246.

function m defined over a Borel field $\mathfrak{A}(E)$ of subsets of E, and suppose that $E \epsilon \mathfrak{A}(E)$ and that $\mathfrak{A}(E)$ is completed with respect to m; i.e., $\mathfrak{A}(E)$ contains all subsets of sets of measure zero. Then the space $L_q(E, m)$ of functions $f(P)$ measurable on E and for which

$$\| f \| = \left[\int_E | f(P) |^q \, dm \right]^{1/q} < \infty$$

is, as is well known, a Banach space whose conjugate space is $L_{q/(q-1)}(E, m)$ in case $q > 1$. For $q = 1$ the conjugate space is the space of essentially bounded and measurable functions with $\| f \| = \text{ess. sup.} \, | f(P) |$.

THEOREM 2. *Suppose that T_α is a uniformly bounded group of linear transformations in $L_q(E, m)$ $(q > 1)$ generated by a group $S(\alpha, P)$ of point transformations of E into itself, i.e.,*

(i) $\quad S(0, P) = P, \qquad S(\alpha + \beta, P) = S(\alpha, S(\beta, P)), \qquad \alpha, \beta \,\epsilon\, E_n\,, P \,\epsilon\, E,$

(ii) $\quad T_\alpha x = x[S(\alpha, \cdot)], \qquad \alpha \,\epsilon\, E_n\,, x \,\epsilon\, L_q(E, m).$

Suppose further that for each x in $L_q(E, m)$ the function $x[S(\alpha, P)]$ is measurable in the product space $E_n \times E$. Then there is a continuous linear projection $y = Ux$ on $L_q(E, m)$ such that

(iii) $\quad T_\alpha Ux = Ux, \qquad \alpha \,\epsilon\, E_n\,, x \,\epsilon\, L_q(E, m),$

(iv) $Ux = \lim_{r \to \infty} V_r^{-1} \int_{I_r} x[S(\alpha, \cdot)] \, d\alpha$, *the limit being in the norm of $L_q(E, m)$.*

(v) *The range $UL_q(E, m)$ of U is the closed linear manifold in $L_q(E, m)$ consisting of those x for which $\alpha \,\epsilon\, E_n$ implies $x[S(\alpha, P)] = x(P)$ for almost all $P \,\epsilon\, E$.*

To prove Theorem 2 it will be sufficient to show that (a) the group is weakly measurable, (b) for each x in $L_q(E, m)$ the function $T_\alpha x$ is almost separably valued, and

(c) $$\int_{I_r} T_\alpha x \, d\alpha = \int_{I_r} x[S(\alpha, \cdot)] \, d\alpha.$$

These three facts, as we shall see, are all immediate consequences of the measurability of $x[S(\alpha, P)]$ in the product space $E_n \times E$. First we shall make another observation. Let x be the characteristic function of the measurable set e. By the group property on $S(\alpha, P)$ we have $x[S(-\alpha, P)] = y(P)$, where y is the characteristic function of $S(\alpha, e)$. Thus, since it was assumed that $T_\alpha x \,\epsilon\, L_q(E, m)$, we see that $S(\alpha, e)$ is measurable if e is measurable. Let x, $\bar{x}$ be the characteristic functions of the measurable sets e, $\bar{e}$, respectively. Then

$$\bar{x} T_\alpha x = \int_e \bar{x}[S(\alpha, P)] \, dm = m[\bar{e} S(-\alpha, e)],$$

and thus by the theorem of Fubini $\bar{x} T_\alpha x$ is measurable. Since the characteristic functions form a fundamental set in L_q as well as in $\bar{L}_q$ (even when $q = 1$),

we see that (a) is satisfied. It might be mentioned at this point that (a) is entirely equivalent to the assertion that $m[\bar{e}S(\alpha, e)]$ is measurable in α for every pair e, $\bar{e}$ of measurable subsets of E. Since E_n is the sum of a denumerable number of cubes, and since the characteristic functions of measurable sets form a fundamental set in $L_q(E, m)$, it will suffice in proving (b) to show that for a characteristic function x, $T_\alpha x$ is almost separably valued when considered as defined only on a cube I. Let $f(\alpha, P) = x[S(\alpha, P)]$, where x is the characteristic function of a measurable set $e \subset E$. Thus f is the characteristic function of a measurable set in the product space $E_n \times E$. There will be therefore a sequence $f_n(\alpha, P)$ of characteristic functions each composed of a finite sum of functions of the form $\phi(\alpha)\psi(P)$, where ϕ, ψ are characteristic functions of measurable sets in E_n, E, respectively, and such that

$$\int_I d\alpha \int_E |f_n(\alpha, P) - f(\alpha, P)|^q \, dm = \int_I d\alpha \int_E |f_n(\alpha, P) - f(\alpha, P)| \, dm \to 0.$$

There is then a sequence $n_i \to \infty$ such that

$$\lim_{i\to\infty} \int_E |f_{n_i}(\alpha, P) - f(\alpha, P)|^q \, dm = 0 \qquad \text{almost everywhere on } I.$$

Thus the function $T_\alpha x = f(\alpha, \cdot)$ has for almost all α in I its values in the closure of the set in $L_q(E, m)$ determined by the set of functional values of $f_n(\alpha, \cdot)$ $(n = 1, 2, \cdots)$. This proves (b). Finally let $\bar{x} = \bar{x}(P)$ be a point in $\overline{L}_q(E, m)$, so that

$$\bar{x} \int_{I_r} T_\alpha x \, d\alpha = \int_{I_r} \bar{x} T_\alpha x \, d\alpha = \int_{I_r} d\alpha \int_E \bar{x}(P) x[S(\alpha, P)] \, dm$$
$$= \int_E \bar{x}(P) \, dm \int_{I_r} x[S(\alpha, P)] \, d\alpha.$$

Then (c) is established.

In order to apply Theorem 1 to the case $\mathfrak{B} = L(E, m)$, we need the following

THEOREM 3. *A set F in $L(E, m)$ is weakly compact if and only if*

$$\text{(i)} \qquad \int_E |f(P)| \, dm \leq C, \qquad f \in F,$$

$$\text{(ii)} \qquad \lim_{m(e)\to 0} \int_e f(P) \, dm = 0 \qquad \textit{uniformly for } f \in F.$$

Any weakly compact set in a Banach space is bounded so that (i) holds if F is weakly compact. A negation of (ii) yields a sequence $e_n \in \mathfrak{A}$ with $m(e_n) \to 0$ and a sequence $f_n \in F$ such that $\left| \int_{e_n} f_n(P) \, dm \right| > \epsilon > 0$. By assumption there is a subsequence $\{g_n\}$ of $\{f_n\}$ such that $\int_e g_n(P) \, dm$ converges for every

$e \in \mathfrak{A}$, and thus by the well-known Hahn-Saks theorem[13] the integrals $\int_e g_n(P)\,dm$ are equi-absolutely continuous. This is the desired contradiction.

To prove the converse we assume (i) and (ii). Let $\{G_i\}$ be a basis for the open sets of real numbers and let $\{f_j\}$ be an arbitrary sequence of points $f_j \in F$. Let E_{ij} be the set in $\mathfrak{A}$ where $f_j(P) \in G_i$, and let $\mathfrak{A}_0$ be the closure in $\mathfrak{A}$ (under the metric $(E', E'') = m(E' - E'') + m(E'' - E')$) of the field determined by the double sequence $\{E_{ij}\}$. Then $\mathfrak{A}_0$ is a Borel field and as a metric space is complete and separable. The space $L_0(E, m)$ of $\mathfrak{A}_0$ measurable functions which are summable on E with respect to m is a closed linear manifold in $L(E, m)$. Let $\{E_j\}$ be dense in the metric space $\mathfrak{A}_0$, and using the diagonal process of G. Cantor, pick a subsequence $\{g_i\}$ of $\{f_i\}$ such that the $\lim_{i\to\infty} \int_{E_j} g_i(P)\,dm$ exists for each $j = 1, 2, \cdots$. Now (ii) implies the uniform (with respect to $f \in F$) continuity of $\int_e f(P)\,dm$ at any point e in $\mathfrak{A}$. Thus we conclude that the $\lim_{i\to\infty} \int_{E_0} g_i(P)\,dm$ exists for every $E_0 \in \mathfrak{A}_0$. This limit, being absolutely continuous with respect to m, determines (by the theorem of Nikodym)[14] an $\mathfrak{A}_0$ measurable function g which is summable with respect to m and such that

$$\lim_{i\to\infty} \int_{e_0} g_i(P)\,dm = \int_{e_0} g(P)\,dm, \qquad e_0 \in \mathfrak{A}_0. \tag{23}$$

Let M_0 be the conjugate of $L_0(E, m)$, i.e., M_0 is the space of bounded $\mathfrak{A}_0$ measurable functions. Let M_0^* be the set of those ϕ in M_0 which assume only a finite number of values. We see by (23) that

$$\lim_{i\to\infty} \int_E \phi(P) g_i(P)\,dm = \int_E \phi(P) g(P)\,dm, \tag{24}$$

for every ϕ in M_0^*. Since M_0^* is dense in M_0 we have, using (i), the fact that (24) holds for every $\phi = M_0$. Now every $\phi \in M = \overline{L}(E, m)$ determines uniquely a $\phi_0 \in M_0$ such that

$$\int_E \phi(P) g(P)\,dm = \int_E \phi_0(P) g(P)\,dm, \qquad g \in L_0(E, m). \tag{25}$$

Thus since g_i, g are in $L_0(E, m)$ we see from (24) and (25) that

$$\int_E \phi(P) g_i(P)\,dm \to \int_E \phi(P) g(P)\,dm, \qquad \phi \in M,$$

which shows that the set F is weakly compact in $L(E, m)$.

[13] S. Saks, *Addition to the note on some functionals*, Trans. Amer. Math. Soc., vol. 35(1933), p. 987.

[14] O. Nikodym, *Sur une généralisation des intégrales de M. J. Radon*, Fund. Math., vol. 15(1930), pp. 131–179.

Simple examples show that (i) is not a consequence of (ii) alone. If, however, E is representable as a sum $\sum_{i=1}^{n(\epsilon)} E_i$ of sets with $m(E_i) < \epsilon$, then (ii) implies (i). This is the case when E is a bounded set in Euclidean space and m is Lebesgue measure.

THEOREM 4. *Let $S(\alpha, P)$ be a group of point transformations satisfying*

(i) $\qquad S(0, P) = P, \qquad S(\alpha + \beta, P) = S(\alpha, S(\beta, P)), \qquad \alpha, \beta \in E_n, P \in E.$

(ii) *If $e \in \mathfrak{A}$, so is $S(\alpha, e) \in \mathfrak{A}$ and $m[S(\alpha, e)] = m[e]$.*

(iii) *For each $x \in L(E, m)$ the function $x[S(\alpha, P)]$ is measurable in the product space $E_n \times E$.*

Then for every $\alpha \in E_n$ and $x \in L(E, m)$ the function $T_\alpha x = x[S(\alpha, \cdot\,)]$ is in $L(E, m)$ and

$$\text{(iv)} \qquad \int_e x[S(\alpha, P)]\, dm = \int_{S(\alpha, e)} x(P)\, dm, \qquad e \in \mathfrak{A},\ \alpha \in E_n,\ x \in L(E, m).$$

$$\text{(v)} \qquad \| T_\alpha x \| = \| x \|.$$

(vi) *There is a continuous linear projection $y = Ux$ of $L(E, m)$ into itself such that*

$$\text{(vii)} \qquad T_\alpha U x = Ux, \qquad x \in L(E, m),\ \alpha \in E_n,$$

(viii) $Ux = \lim_{r \to \infty} V_r^{-1} \int_{I_r} x[S(\alpha, \cdot\,)]\, d\alpha$, *the limit being in the norm of $L(E, m)$, i.e.,*

$$\lim_{r \to \infty} \int_E \left| y(P) - \frac{1}{V_r} \int_{I_r} x[S(\alpha, P)]\, d\alpha \right| dm = 0, \qquad y = Ux.$$

(ix) *The range $UL(E, m)$ of U is the closed linear manifold in $L(E, m)$ consisting of those x for which $\alpha \in E_n$ implies $x[S(\alpha, P)] = x(P)$ for almost all $P \in E$.*

An elementary calculation based on (i) and (ii) shows that (iv) and (v) hold for $\mathfrak{A}$ measurable functions which assume only a finite number of values. If $\{x_q\}$ is a sequence of such functions approaching x in the mean as well as almost everywhere, then $x_q[S(\alpha, P)] \to x[S(\alpha, P)]$ almost everywhere and the integrals

$$\int_e x_q[S(\alpha, P)]\, dm = \int_{S(\alpha, e)} x_q(P)\, dm$$

are equi-absolutely continuous. Thus $x[S(\alpha, P)]$ is summable in P and equations (iv) and (v) hold for every x in $L(E, m)$. The conditions (iv) and (v) show, when combined with Theorem 3, that for a given $x \in L(E, m)$ the set $T_\alpha x$ $(\alpha \in E_n)$ is weakly compact. It then follows immediately that the set $V_r^{-1} \int_{I_r} T_\alpha x\, d\alpha$ is weakly compact; for

$$\int_e dm \left[V_r^{-1} \int_{I_r} x[S(\alpha, P)]\, d\alpha \right] = V_r^{-1} \int_{I_r} d\alpha \int_e x[S(\alpha, P)]\, dm$$

$$= V_r^{-1} \int_{I_r} d\alpha \int_{S(\alpha, e)} x(P)\, dm \leqq \sup_{\alpha} \int_{S(\alpha, e)} x(P)\, dm.$$

Thus Theorem 4 follows from Theorem 1.

THEOREM 5. *Let the group* T_α $(\alpha \in E_n)$ *satisfy the conditions of Theorem* 1. *Then for each* x *in* $\mathfrak{B}$ *the function* $T_\alpha x$ *is uniformly continuous on* E_n.

In the case of semi-groups on E_n^+ we can only state a sort of one-sided continuity in the interior of E_n^+, i.e., $T_\alpha x \to T_\beta x$ if $\alpha \to \beta$ in such a way that its components are greater than those of β. Theorem 5 may be proved by the same method we have used for one parameter groups[15] and so we omit the details. The uniform continuity stated here follows from the uniform boundedness of the group.

YALE UNIVERSITY.

[15] N. Dunford, *On one parameter groups of linear transformations*, Annals of Math., vol. 39(1938), pp. 569–573.

THE EXISTENCE OF CERTAIN TRANSFORMATIONS

By G. T. Whyburn

The general existence problem for a specified type of transformation of one given compact set onto another is of long standing and has received numerous contributions over a considerable span of years. For example, a classical result of Hahn and Mazurkiewicz yields the conclusion that any compact locally connected continuum can be mapped continuously onto any other one but cannot be so mapped onto a non-locally-connected continuum.

In this paper the question of the mappability of a compact locally connected continuum M onto an interval by particular sorts of continuous transformations will be considered. This is, of course, the same thing as considering the definability of particular kinds of continuous, real-valued functions on M. In this connection the reader is referred to closely related papers by Čech,[1] Mazurkiewicz,[2] Aitchison,[3] C. Pauc,[4] Kuratowski,[5] and the author.[6]

We consider monotone, non-alternating, interior and light transformations. If A and B are compact continua, a continuous transformation $T(A) = B$ is (1) *monotone*[7] provided the inverse set $T^{-1}(b)$ of each point b in B is connected, (2) *non-alternating*[8] if for any two points x and y of B, $T^{-1}(x)$ does not separate any two points of $T^{-1}(y)$ in A, (3) *interior*[9] provided the image of every set open in A is open in B, and (4) *light* provided that for each point b in B, $T^{-1}(b)$ is totally disconnected (or of dimension 0).

The principal results will be found in §§2 and 4. In §2 it is shown that a compact locally connected continuum M can be mapped onto an interval by a non-alternating interior transformation f if and only if the cyclic elements of M are arranged into a cyclic chain. In §4 it is shown that in case M is 1-dimensional, f can in addition be chosen as a light transformation.

1. **Lemmas on joining and subdivision.** If a and b are points of a locally connected continuum M and K is the set of all points separating a and b in M,

Received March 22, 1939.

[1] Fundamenta Mathematicae, vol. 18(1932), p. 85.

[2] Ibid., p. 88.

[3] Comptes Rendus de la Société des Sciences et des Lettres de Varsovie, vol. 27(1934).

[4] Comptes Rendus, Paris, vol. 202(1939), p. 489.

[5] Fundamenta Mathematicae, vol. 30(1938), p. 17.

[6] American Journal of Mathematics, vol. 55(1933), p. 131.

[7] See R. L. Moore, Transactions of the American Mathematical Society, vol. 27(1925), p. 416; C. B. Morrey, American Journal of Mathematics, vol. 57(1935), p. 17; also reference in footnote 8.

[8] See G. T. Whyburn, American Journal of Mathematics, vol. 56(1934), pp. 394–402.

[9] See Stoïlow, Annales Scientifiques de l'Ecole Normale Supérieure, vol. 63(1928), pp. 347–382.

the *cyclic chain* $C(a, b)$ in M from a to b is the set consisting of $K + a + b$ plus all cyclic elements of M intersecting $K + a + b$ in exactly two points. The set $C(a, b)$ is known also to consist of all simple arcs in M of the form axb. In this paper we shall have occasion frequently to impose the condition that M be identical with some cyclic chain $C(a, b)$ in M.

(1.1) LEMMA. *Let M be a compact locally connected continuum such that $M = C(a, b)$ for $a, b \in M$. Let $A \supset a$, $B \supset b$ be continua in M not separating M. Then for any $x \in M-(A + B)$, there exists an irreducible continuum AxB in M between A and B containing x which is locally an arc at all points of $AxB-(A + B)$.*

Proof. Let G be the upper semi-continuous decomposition of M into the sets A, B and points of $M-(A + B)$. Let $T(M) = M'$ be the associated (monotone) transformation. Let $T(A) = a'$, $T(B) = b'$. Then M' is a locally connected continuum; and since a' and b' are non-cut points of M' and T is monotone, it readily follows that $M' = C(a', b')$. Hence M' contains an arc $a'xb'$ and we have only to set

$$AxB = T^{-1}(a'xb').$$

Let the locally connected continuum M be a cyclic chain $C(a, b)$. By a *subdivision* σ of M will be meant a finite, linearly ordered, set of disjoint irreducible cuttings of M between a and b.

$$a = X_0, X_1, X_2, X_3, \cdots, X_n, X_{n+1} = b,$$

where X_i $(1 \leqq i \leqq n)$ cuts M into two connected sets $M_a(X_i)$, $M_b(X_i)$ so that $M_a(X_i) \supset a + \sum_{j=1}^{i-1} X_j$ and $M_b(X_i) \supset b + \sum_{i+1}^{n} X_j$. The set

$$X_{i-1} + X_i + M_a(X_i) \cdot M_b(X_{i-1}) = I_i$$

will be called the *interval* from X_{i-1} to X_i $(a = X_0, b = X_{n+1})$. A set of the form $M_a(X_i) \cdot M_b(X_{i-1})$ will be called an *open interval* of σ.

(1.2) LEMMA. *Given any subdivision σ_0 and any $\epsilon > 0$, there exists a subdivision σ containing σ_0 and such that if X_i is any element of σ,*

$$V_\epsilon(X_i) \supset I_i + I_{i+1}.$$

Proof. For convenience we will suppose the metric in M is the so-called "relative distance" of Mazurkiewicz.[10] Now let $e = \frac{1}{8}\epsilon$. Since the sets $V_e(x)$, $x \in M$, are open (and connected) and cover M, we can select a finite number of them

$$V_1 = V_e(x_1), \quad V_2 = V_e(x_2), \quad \cdots, \quad V_m = V_e(x_m)$$

which cover M.

[10] Fundamenta Mathematicae, vol. 1(1920), p. 167. In this metric, spherical neighborhoods of points or of connected sets are connected. Note: For any set X and any $r > 0$, $V_r(X)$ denotes the set of all points at a distance $< r$ from X.

Now consider V_1. If $\bar{V}_1 \supset b$, we need go no further. Suppose $\bar{V}_1$ does not contain b. Let X_{i+1}^0 be the first element of σ_0 such that $M_a(X_{i+1}^0) \supset \bar{V}_1$. Now if $\overline{V_{2e}(x_1)} \cdot X_{i+1}^0 \neq 0$, we need go no further. If this is not so we proceed as follows. Applying (1.1), using the continua $\alpha = M_a(X_i^0) + X_i^0$ and $\beta = M_b(X_{i+1}^0) + X_{i+1}^0$, we obtain a continuum $\alpha x \beta$ which is locally an arc at x, where $x \in V_1$. Let y be the last point of $\overline{V_{2e}(x_1)}$ on $\alpha x \beta$ in the order α, β, and set

$$A = \bar{V}_1 + \alpha x \text{ (of } \alpha x \beta), \qquad B = y\beta \text{ (of } \alpha x \beta).$$

Then A and B are disjoint continua; accordingly[11] there exists a set X' which separates M irreducibly between A and B and indeed such that $M - X'$ has just two components each bounded by X'. Clearly X' is on the "b side" of V_1 (i.e., $M_a(X') \supset V_1$), and $X' \cdot V_{2e}(x_1) \neq 0$ since $A \cdot V_{2e}(x_1) \neq 0 \neq B \cdot \overline{V_{2e}(x_1)}$. In exactly the same manner we construct a set X'' on the "a side" of V_1 such that $V_{2e}(x_1) \cdot X'' \neq 0$. Now let us add X' and X'' to σ_0 and call σ_1 the resulting subdivision.

Next let us proceed with V_2 and σ_1 in exactly the same way as we did above with V_1 and σ_0 and obtain a subdivision σ_2 containing elements X' and X'' each intersecting $V_{2e}(x_2)$ and such that V_2 is in the interval of σ_2 from X'' to X', i.e.,

$$V_2 \subset M_b(X'') \cdot M_a(X').$$

Continuing in this manner to V_m, we obtain a subdivision σ_m, which we call σ, having the property that for any $i \leqq m$ there are elements X_j, X_k of σ intersecting $V_{2e}(x_i)$ and such that

$$V_i \subset M_b(X_j) \cdot M_a(X_k).$$

Since $e = \frac{1}{5}\epsilon$ and the sets $[V_i]$ cover M, clearly this is equivalent to our lemma.

2. Non-alternating interior transformations into an interval.

(2.1) THEOREM. *If the locally connected continuum M is a cyclic chain $C(a, b)$, there exists a non-alternating interior mapping f of M onto the interval* (0, 1) *such that $f^{-1}(0) = a$, $f^{-1}(1) = b$.*

Proof. By (1.2) we can set up an infinite monotone sequence of subdivisions $\sigma_1, \sigma_2, \cdots$ of M such that for each n, σ_n satisfies the conclusions of (1.2) for $\epsilon = n^{-1}$.

Let G denote the collection of all the elements in all of the subdivisions σ_i. Then G is a non-separated collection (see footnote 11) of cuttings of M between a and b. Furthermore, if $X \in G$ and $p \in M - X$, by Lemma (1.2) there exists a σ_n containing X and such that neither interval of σ_n abutting on X contains p. Thus there exists an element Y of G (in fact of σ_n) which separates X and p in M.

[11] See the author's paper in the Transactions of the American Mathematical Society, vol. 33(1931), pp. 444–454. A collection G of disjoint subsets of a connected set M is called *non-separated* provided no element of G separates in M two points lying on another element of G.

Hence by a result previously established,[12] if for each $z \in M - G'$ $(G' = \sum X)$ we let Z be z together with all $p \in M$ which cannot be separated in M from z by any single element of G, then Z is closed; and if we call G_0 the collection obtained by adding to G all such sets Z, then G is upper semicontinuous. Furthermore, every element of G_0 except the ones containing a and b must separate a and b in M. For clearly each new element Z can be represented in the form

$$Z = \prod_1^\infty I_n,$$

where I_n is the interval of σ_n containing z.

Hence the hyperspace of G_0 is a simple arc which we may suppose is the interval $(0, 1)$. Let f be the associated transformation of M onto $(0, 1)$. Since[12] the collection G_0 is non-separated, it follows that f is non-alternating. To see that f is also interior, let U be any open set in M and let $x \in U$. There exists a region R, $x \subset R \subset U$. If $x \in X \in G$, then since R intersects each of the regions $M_a(X)$ and $M_b(X)$ which map into the open intervals $[a, f(X)]$ and $[f(X), b]$ respectively, it follows that $f(R)$ and hence $f(U)$ contains an open interval about $f(X) = f(x)$. If x belongs to no element of G, it follows by the way Lemma (1.2) gave the sets of G that there exist elements X_j and X_k of a σ_n such that $X_j \cdot R \neq 0 \neq X_k \cdot R$ and x belongs to the interval from X_j to X_k. But this gives $f(R) \supset [f(X_j), f(X_k)] \supset f(x)$, so that $f(U)$ contains an open interval about $f(x)$, and our proof is complete.

It will be noted that the transformation f is so defined that for any x with $0 < x < 1$, $f^{-1}(x)$ separates M irreducibly between a and b into exactly two regions. Thus we have

(2.11) *If M is unicoherent, f will be monotone.*

The property of f just mentioned actually is a necessary consequence of the properties sought in Theorem (2.1) as will now be shown.

(2.2) THEOREM. *Let $f(x)$ be a non-alternating interior mapping of a locally connected continuum M onto the interval* $(0, 1)$ *and let* $A = f^{-1}(0)$, $B = f^{-1}(1)$. *Then for each y with* $0 < y < 1$, $f^{-1}(y)$ *separates M irreducibly between A and B into just two components. Neither A nor B separates M and if $a \in A$, $b \in B$, M is identical with the cyclic chain $C(a, b)$.*

Proof. By known properties of non-alternating transformations (see footnote 8) it follows that $f^{-1}(y)$ separates M into just two components R_a and R_b containing A and B, respectively. Furthermore R_a maps onto $(0, y) - y$ and R_b onto $(y, 1) - y$. Thus since f is interior, $f^{-1}(y)$ must be the boundary both of R_a and of R_b. Hence $f^{-1}(y)$ separates M irreducibly.

Since neither 0 nor 1 separates $(0, 1)$ and f is non-alternating, it follows (see footnote 8) that neither A nor B separates M.

[12] Ibid., pp. 451–452. The equivalence used here of the property of a transformation being non-alternating and the property of its associated decomposition being non-separated is perhaps worthy of special note.

To show that $M = C(a, b)$, let us suppose on the contrary that there is a component R of $M - C(a, b)$. The boundary of R is a single point x and x cannot belong to A or to B since neither of these sets separates M. Thus $x \in f^{-1}(y)$ for some y with $0 < y < 1$. But this is impossible since $f^{-1}(y)$ would separate M into at least three components. Thus $M = C(a, b)$, and our proof is complete.

Combining (2.1) and (2.2), we have

(2.3) THEOREM. *In order that a compact locally connected continuum M be mappable onto the interval $(0, 1)$ by a non-alternating interior transformation f so that $f(a) = 0$, $f(b) = 1$, $a, b \in M$, it is necessary and sufficient that M be identical with the cyclic chain $C(a, b)$.*

3. Lemmas on separation.

(3.1) LEMMA. *Let M be a locally connected continuum, let $f(M) = (0, 1)$ be continuous and $a \in f^{-1}(0)$, $b \in f^{-1}(1)$. For each x, $0 < x < 1$, we can choose a subset X of $f^{-1}(x)$ such that the collection $[X]$ is a non-separated collection of cuttings of M between a and b.*

Proof. For each x, $0 < x < 1$, let $R_a(x)$ be the component of $M - f^{-1}(x)$ containing a, let $R_b(x)$ be the component of $M - \overline{R_a(x)}$ containing b and let $X = F[R_b(x)]$.[13] Then clearly

$$X = F[R_a(x)] \cdot F[R_b(x)] \subset f^{-1}(x).$$

Further, if $0 < x_1 < x_2 < 1$, we have

$$X_1 \subset \overline{R_a(x_1)} \subset R_a(x_2), \qquad X_2 \subset \overline{R_b(x_2)} \subset R_b(x_1).$$

Hence the collection $[X]$ is non-separated.[14]

(3.2) LEMMA. *If a and b are points of a 1-dimensional locally connected continuum M and F is any closed 0-dimensional subset of M, there exists a 0-dimensional set $X \subset M - F$ which separates M irreducibly between a and b into just two components.*

Proof. There exists (see footnote 6) a light transformation $f(x)$ of M onto $(0, 1)$ so that $f(a) = 0$, $f(b) = 1$ and $f(F) \subset f(a) + f(b)$. Hence, applying (3.1), we obtain an uncountable, non-separated collection of 0-dimensional cuttings X of M between a and b no one of which intersects F. Thus[15] some one of these sets (in fact all but a countable number of them) separates M irreducibly between a and b into just two components.

The results just established lead at once to

(3.3) THEOREM. *If A and B are disjoint subcontinua of a 1-dimensional locally connected continuum M and F is any 0-dimensional closed subset of M,*

[13] For any open set R, $F(R)$ denotes the boundary of R, i.e., the set $\bar{R} - R$.

[14] Compare this lemma and its proof with that given by the author on pp. 452–453 of the paper cited in footnote 11.

[15] See p. 450 of the paper cited in footnote 11.

there exists a 0-dimensional set $X \subset M - F$ which separates M irreducibly between A and B into just two components.

Proof. Let α denote the continuum obtained by adding to A all components of $M - A$ except the one containing B; and similarly, let $\beta = B$ plus all components of $M - B$ except the one containing A. Let M' be the hyperspace of the decomposition of M into the sets α, β and individual points of $M - (\alpha + \beta)$ and let $T(M) = M'$ be the associated monotone transformation. Let $T(\alpha) = a$, $T(\beta) = b$. Then M' is 1-dimensional and locally connected; and, applying (3.2) to M', we get a 0-dimensional set $X' \subset M' - T(F)$ which separates M' irreducibly between a and b into just two components. Then if we set $X = T^{-1}(X')$, it follows by known properties of monotone transformations that X satisfies our theorem.

4. **Non-alternating interior light transformations into an interval.** Let M be a 1-dimensional locally connected continuum which is identical with a cyclic chain $C(a, b)$. A subdivision σ of M will be called 0-dimensional if each element of it is 0-dimensional. A subdivision σ' will be called a *refinement* of σ provided it is obtained from σ by inserting additional elements. It will be understood that any refinement of a 0-dimensional subdivision is itself 0-dimensional. Actually in this section we shall deal only with 0-dimensional subdivisions.

(4.1) LEMMA. *Given a 0-dimensional subdivision σ of M and any $\epsilon > 0$, there exists a refinement σ' of σ such that each component of an open interval in σ' is of diameter $< \epsilon$.*

Proof. Let K be the sum of all elements of σ and let $d = \frac{1}{10}\epsilon$. By well-known decomposition theorems for 1-dimensional locally connected continua,[16] there exists a 0-dimensional closed set F in $M - K$ and a decomposition $M = \sum_1^n M_i$ such that for each i, M_i is a locally connected continuum of diameter $< d$ and for $i \neq j$, $M_i \cdot M_j \subset F$.

It follows at once from (3.1) and the argument given under (1.2) that there exists a refinement σ_1 of σ such that if K_1 denotes the sum of all elements of σ_1, then $K_1 \cdot F = 0$ and for each $i \leqq n$, $K_1 \cdot M_i \neq 0$.

Now let M_i and M_j be any two of the sets $[M_i]$ such that $\rho(M_i, M_j) > \frac{1}{2}\epsilon$. It will be shown that

(*) *there exists a refinement σ_k of σ_1 such that no component of an open interval of σ_k intersects both M_i and M_j.*

To this end, let $I = XY$ be any closed interval of σ_1 such that the open interval $E = I - (X + Y)$ intersects both M_i and M_j. Let F_i and F_j be the boundaries of M_i and M_j, respectively; i.e., $F_i = F \cdot M_i$, $F_j = F \cdot M_j$. Since the closure of any component of $M_i \cdot E$ (or $M_j \cdot E$) must intersect either both X and Y or both $X + Y$ and F_i (or F_j) and $K_1 \cdot F = 0$, it follows that $M_i \cdot E$ and $M_j \cdot E$ have finite numbers, say n_i and n_j respectively, of components.

[16] See, for example, Menger's *Kurventheorie*, p. 191. Note also references there given to Urysohn and Vanek.

Now for any component R of $M_i \cdot E$ whose closure intersects both X and Y, let V be a neighborhood of Y so that $V \cdot F_i = 0$. Then R contains a connected subset S such that $\bar{S} \cdot Y = 0$ and $S \supset R \cdot (M - V)$. Thus any component of E which intersects R but is not contained wholly in R must intersect S. Let us replace each such component R of $M_i \cdot E$ by the corresponding set S and call S_i the resulting set. Similarly let us obtain a set S_j in $M_j \cdot E$. Then S_i and S_j are subsets of M_i and M_j having n_i and n_j components respectively and such that (1) any component of E which intersects both M_i and M_j must intersect both S_i and S_j and (2) the closure of each component of S_i or of S_j intersects exactly one of the sets X and Y.

Let us add to $M_a(X) + X$ the closure of all components of S_i and of S_j whose closures intersect X and call A the continuum thus obtained. Similarly, let $B = M_b(Y) + Y$ plus all components of S_i and of S_j whose closures intersect Y. Applying (3.3) to A and B, we obtain a 0-dimensional set Z irreducibly separating M between A and B into just two components and not intersecting F.

Now there are, say, m_i components of S_i and m_j components of S_j in the interval XZ and the closure of each of these intersects X but not Z; and there are $n_i - m_i$ components of S_i and $n_j - m_j$ components of S_j in ZY and the closure of each of these intersects Y but not Z.

Consider the interval XZ. Since $M_b(X)$ is connected, there is some component P of S_i or of S_j, say of S_i, which can be joined by an arc α in $M_b(X)$ to a point of Z such that α intersects no other component of S_i or of S_j. Now, just as we replaced R by S above, we can replace P by a connected subset Q of P with $\bar{Q} \cdot \bar{X} = 0$ which also intersects α and has the property that any component of $XZ - (X + Z)$ which intersects both S_j and P also intersects Q. Let B_1 denote the continuum $M_b(Z) + Z + \alpha + \bar{Q}$ and let A_1 be the continuum $M_a(X) + X + \overline{S_i - P} \cdot XZ + \bar{S}_j \cdot XZ$. Applying (3.3) to A_1 and B_1, we obtain a 0-dimensional set W_1 irreducibly separating M between A_1 and B_1 into just two components and not intersecting F. Further, the interval XW_1 has the property that any component of XW_1 which intersects both M_i and M_j must intersect one of the $m_i - 1$ components of S_i which are in XW_1, and the closure of each of these components as well as of each component of S_j in XW_1 intersects X but not W_1. Repeating this argument, we get a set W_2 between X and W_1 which separates off either a second component of S_i in XW_1 or one of S_j in XW_1, and so on. After at most $m_i + m_j$ steps we obtain a subdivision $X, W_e, W_{e-1}, \cdots, W_2, W_1, Z$ of XZ so that no component of an open interval in this subdivision can intersect both M_i and M_j.

In exactly the same manner we can subdivide the interval ZY. Hence we have shown that each interval XY in σ_1 can be subdivided so that no component of an open interval in the refinement can intersect both M_i and M_j. Let us so subdivide each interval in σ_1 and call σ_k the resulting refinement of σ_1. Then (*) is satisfied.

Now applying (*) successively to all pairs M_i, M_j satisfying $\rho(M_i, M_j) > \frac{1}{2}\epsilon$, we obtain eventually a refinement σ' of σ which must satisfy our lemma. For if there were a component R of an open interval of σ' of diameter $> \epsilon$, we could

find points $p, q \in R$ with $\rho(p, q) > \frac{3}{4}\epsilon > \frac{1}{2}\epsilon + 2d$. Hence if $M_i \supset p, M_j \supset q$, we have $\rho(M_i, M_j) > \rho(p, q) - 2d > \frac{1}{2}\epsilon$. Accordingly no component of an open interval of σ' could intersect both M_i and M_j, and we have a contradiction.

(4.2) THEOREM. *If the 1-dimensional locally connected continuum M is identical with its cyclic chain $C(a, b)$, $a, b \in M$, there exists a non-alternating light interior transformation $f(x)$ of M onto the interval* (0, 1) *so that* $f^{-1}(0) = a, f^{-1}(1) = b$.

Proof. We note first that by virtue of (3.3), under the conditions of our theorem, we can obtain (1.2) with the stronger conclusion obtained by substituting "0-dimensional subdivision σ" for the words "subdivision σ".

Thus by using (4.1) in addition, we can set up a monotone sequence of 0-dimensional subdivision $\sigma_1, \sigma_2, \cdots$ of M such that, for each n, σ_n satisfies the conclusion of both (1.2) and (4.1) for $\epsilon = n^{-1}$.

Now by exactly the argument given in the proof of (2.1), beginning with the second paragraph, we define our function $f(x)$ and show that it satisfies all our conclusions except the property of being light. But that f is light results at once from (4.1) and the fact that each σ_n is 0-dimensional. For if $0 < y < 1$, the set $f^{-1}(y)$ either belongs to σ_n for some n (which gives $\dim f^{-1}(y) = 0$) or else it is of the form

$$f^{-1}(y) = \prod_{1}^{\infty} E_n,$$

where E_n is an open interval of σ_n. In the latter case, since each component of E_n is of diameter $< n^{-1}$, $f^{-1}(y)$ must be 0-dimensional.

The theorem just proved together with a known result[17] on light interior transformations gives the following interesting additional conclusion.

For any point x of M there exists a simple arc axb in M which maps topologically onto the interval (0, 1) *under f.*

Thus $f(x)$ gives a sort of "arc development"[18] to M under the conditions of (4.2).

5. **Interior transformations of graphs and dendrites onto an interval.** It is known[17] that any interior transformation on a graph or a dendrite is necessarily light. Also it follows from (2.2) that the only dendrite which can be mapped onto an interval by a non-alternating interior transformation is the simple arc. However, if we omit the non-alternating requirement on the transformation, we find that a more inclusive but not all-inclusive class of dendrites can be mapped onto an interval. We begin with

(5.1) *Any connected graph A can be mapped interiorly onto an interval.*

For let X be the set of vertices in a subdivision of A and let Y be a set obtained by selecting one point from each component of $A - X$. Then if we define

[17] See my paper in this Journal, vol. 3(1937). Note also references there given to Stoïlow and Montgomery.

[18] Compare this with results recently announced by C. Pauc. See footnote 4.

$$f(x) = 0, \qquad \text{if } x \in X;$$
$$f(x) = 1, \qquad \text{if } x \in Y$$

and let f map the closure of each component of $A-(X+Y)$ topologically onto $(0, 1)$ preserving end points, clearly f maps A interiorly onto $(0, 1)$.

(5.2) THEOREM. *In order that a dendrite A be mappable interiorly onto an interval it is necessary and sufficient that* (a) *no point of A cut A into infinitely many components, and* (b) *if H is the set of all end points of A, $\bar{H} - H$ be a finite set.*

Proof. To show the conditions necessary, let $f(A) = (0, 1)$ be interior and let $K = f^{-1}(0) + f^{-1}(1)$. Then since every end point of A must (see footnote 17) map into an end point of $(0,1)$, we have $K \supset H$; and by a previous result of the author[19] it follows that $K - H$ is a finite set. Then since K is closed, it results that (b) is necessary. Now suppose, contrary to (a), that a point x of A cuts A into infinitely many components. Then since x is a limit point of H, we have $x \in K$, i.e., $f(x) = 0$ or $f(x) = 1$; but since there is only one component of $(0, 1) - f(x)$, there can be only a finite number of components of $A - f^{-1}f(x)$. Clearly this is a contradiction.

To show the conditions sufficient, let us suppose they are satisfied in A. Since $A-H$ is connected and $\bar{H}-H$ is finite and (a) is satisfied, it follows that $A-\bar{H}$ has just a finite number, say n, of components. Now let N be a set containing just one point of order 2 from each component of $A-\bar{H}$. Then $A-\bar{H}-N$ has just $2n$ components $R_1, R_2, \cdots, R_{2n}$ each of which has just one limit point in N and one or more limit points in $\bar{H}$.

Let R be any one of the sets R_i, let $x = \bar{R}\cdot N$ and $Y = \bar{R}\cdot\bar{H}$. Decompose $\bar{R}$ into the set Y and the individual points of $\bar{R} - Y$, let R' be the hyperspace of this decomposition and let $T_1(\bar{R}) = R'$ be the associated transformation. Let $T(x) = x'$, $T(Y) = y'$. Since there can be no end point of $\bar{R}$ in R, it readily follows that R' is identical with the cyclic chain $C(x', y')$ in R'. Hence by (2.1) there exists a non-alternating interior transformation $T_2(R') = (0, 1)$ such that $T_2^{-1}(0) = x$, $T_2^{-1}(1) = y'$. It readily follows that the transformation T_2T_1 is interior.

Let us now define $f(x)$ as follows:

$$f(x) = \begin{cases} 0 & \text{for } x \in N; \\ 1 & \text{for } x \in \bar{H}; \\ T_2^iT_1^i(x) & \text{for } x \in R_i, \end{cases}$$

where T_1^i and T_2^i are defined for R_i as T_1 and T_2 were defined above for R. Then since each $T_2^iT_1^i$ is interior, and since any open set in A intersecting N or $\bar{H}$ contains a connected open subset of some R_i which maps into a non-degenerate interval abutting on 0 or 1 respectively, it follows that $f(x)$ is interior.

THE UNIVERSITY OF VIRGINIA.

[19] See this Journal, vol. 4(1938), p. 609.

THE WEIERSTRASS CONDITION FOR MULTIPLE INTEGRAL VARIATION PROBLEMS

By Lawrence M. Graves

Consider the problem of minimizing a multiple integral

$$I = \int_R f(t, x, p)\, dt = \int \cdots \int f(t, x, p)\, dt_1 \cdots dt_n$$

in a class of admissible manifolds in (t, x)-space with equations in the form $x = x(t)$, where $x = (x^1, \cdots, x^m)$, $t = (t_1, \cdots, t_n)$, and p denotes the matrix $(p^i_\alpha) = (\partial x^i/\partial t_\alpha)$. The Weierstrass E-function has the form

$$E(t, x, p, P) = f(t, x, P) - f(t, x, p) - (P^i_\alpha - p^i_\alpha) f^\alpha_i(t, x, p),$$

where $f^\alpha_i = \partial f/\partial p^i_\alpha$ and the usual summation convention is used. We suppose that f and its partial derivatives f^α_i are continuous in a certain region S of (t, x)-space for all p.[1] For the class of admissible manifolds we take all manifolds $x = x(t)$ lying in S, of class D′ on the fixed region R of t-space, and having a fixed boundary over the boundary of R. Then *a necessary condition for a minimum of I is that $E(t, x, p, P) \geqq 0$ for all (t, x, p) on the minimizing manifold and for all P such that the matrix $P - p = (P^i_\alpha - p^i_\alpha)$ has rank one.*

A very brief and elementary proof for the condition is given in §1 below.[2] The proof is similar to one given by the author for simple integral problems.[3] The Weierstrass condition as stated obviously applies also to problems in parametric form. In §2 the condition is transformed so as to be expressed entirely in terms of the n-rowed minors of the matrices p and P.[4]

1. **Proof of the Weierstrass condition.** Let the minimizing manifold M_0 have equations $x^i = \phi^i(t)$, and suppose that the partial derivatives $\partial\phi^i/\partial t_\alpha$ are continuous near $t = \bar{t}$. Let these derivatives for $t = \bar{t}$ be denoted by p^i_α. If the matrix $\Delta p = P - p$ has rank one, it may be represented in the form $\Delta p^i_\alpha =$

Received March 27, 1939.

[1] Obviously we could also consider restrictions on the admissible values of the p^i_α.

[2] For more general classes of admissible manifolds McShane has shown that $E \geqq 0$ almost everywhere. See Annals of Mathematics, vol. 32(1931), pp. 578–590. Another type of proof has been developed by Coral. See this Journal, vol. 3(1937), pp. 585–592.

[3] *A proof of the Weierstrass condition in the calculus of variations*, American Mathematical Monthly, vol. 41(1934), pp. 502–504.

[4] In the cases $n = 1$ and $n = m - 1$ (with more general hypotheses on the class of admissible manifolds) McShane gave a direct proof of the transformed condition for parametric problems. See Annals of Mathematics, vol. 32(1931), pp. 723–733.

$P^i_\alpha - p^i_\alpha = F_\alpha G^i$, where $F_\alpha F_\alpha > 0$, $G^i G^i > 0$. Then there exists a matrix $(B_{\gamma\alpha})$ such that

$$\begin{pmatrix} B_{\gamma\alpha} \\ F_\alpha \end{pmatrix} \qquad (\gamma = 1, \cdots, n-1; \alpha = 1, \cdots, n)$$

is non-singular. Introduce the new variables $u_\gamma = B_{\gamma\alpha}(t_\alpha - \bar{t}_\alpha)$, $v = F_\alpha(t_\alpha - \bar{t}_\alpha)$. Let $\eta(u, b)$ be a function which vanishes for $u_\gamma u_\gamma = b^2$, is positive for $u_\gamma u_\gamma < b^2$, and is such that η and $\partial\eta/\partial u_\gamma$ approach zero with b. For example, we may take $\eta = (1 - u_\gamma u_\gamma)^{\frac{1}{2}} - (1 - b^2)^{\frac{1}{2}}$. Let $C^i = (\partial\phi^i/\partial v)^{t=\bar{t}} + G^i$, and set

$$X^i(u, v) = \phi^i(u, 0) + C^i v. \tag{1}$$

Then at $t = \bar{t}$ we find

$$\frac{\partial X^i}{\partial t_\alpha} = \frac{\partial \phi^i}{\partial t_\alpha} + F_\alpha G^i = P^i_\alpha. \tag{2}$$

Suppose $0 < \epsilon < 1$, and for $u_\gamma u_\gamma < b^2$ set

$$\omega^i(u, v; b) = \phi^i(u, \eta) + (v - \eta)\,\frac{[\phi^i(u, \eta) - X^i(u, \epsilon\eta)]}{\eta - \epsilon\eta}.$$

By use of (1) and the assumption that $\partial\eta/\partial u_\gamma$ tends to zero with b, we find that if the partial derivatives of ω^i are evaluated for $u_\gamma u_\gamma < b^2$, $0 < v < \eta$, then as b tends to zero,

$$\frac{\partial \omega^i}{\partial u_\gamma} \to \left(\frac{\partial \phi^i}{\partial u_\gamma}\right)^{t=\bar{t}},$$

$$\frac{\partial \omega^i}{\partial v} \to \frac{1}{1-\epsilon}\left[\left(\frac{\partial \phi^i}{\partial v}\right)^{t=\bar{t}} - \epsilon C^i\right] = \left(\frac{\partial \phi^i}{\partial v}\right)^{t=\bar{t}} - \frac{\epsilon}{1-\epsilon}\, G^i.$$

Consequently

$$\frac{\partial \omega^i}{\partial t_\alpha} \to \left(\frac{\partial \phi_i}{\partial t_\alpha}\right)^{t=\bar{t}} - \frac{\epsilon}{1-\epsilon}\, F_\alpha G^i = p^i_\alpha - \frac{\epsilon}{1-\epsilon}\, \Delta p^i_\alpha. \tag{3}$$

Now let R_1 denote the region of t-space determined by the inequalities $u_\gamma u_\gamma < b^2$, $0 < v \leqq \epsilon\eta$, and let R_2 denote the region determined by $u_\gamma u_\gamma < b^2$, $\epsilon\eta < v < \eta$. If $W(b)$ is the n-dimensional "volume" of $R_1 + R_2$, then $\epsilon W(b)$ is the volume of R_1 and $(1 - \epsilon)W(b)$ is the volume of R_2. Let the manifold M_b be defined as follows:

$$x^i = \begin{cases} X^i(t) & \text{on } R_1, \\ \omega^i(t; b) & \text{on } R_2, \\ \phi^i(t) & \text{elsewhere.} \end{cases}$$

Then M_b coincides with the minimizing manifold M_0 except over the region $R_1 + R_2$ and is an admissible manifold when b is sufficiently small, except possibly when the portion of M_0 over $R_1 + R_2$ has points in common with the boundary

of the region S where admissible manifolds must lie. Let us assume that the boundary D of the region S is a manifold of class C'. Then in case M_b turns out to lie on the wrong side of D, a change of the signs of the F_α and G^i leads to a manifold M_b on the right side of D, except possibly when the manifold $x^i = X^i(t)$ is tangent to D. In this exceptional case a slight modification of the matrix P leads to a manifold $x^i = X^i(t)$ which is not tangent to D, and then considerations of continuity show that the desired result still holds. Thus, since our method of proof makes no use of the vanishing of the first variation, it is applicable to the case of "unilateral variations".

Since $I(M_0)$ is a minimum we have

$$0 \leqq \frac{I(M_b) - I(M_0)}{W(b)} = \frac{\epsilon}{\epsilon W(b)} \int_{R_1} f(t, X, X_t)\, dt + \frac{1-\epsilon}{(1-\epsilon)W(b)} \int_{R_2} f(t, \omega, \omega_t)\, dt - \frac{1}{W(b)} \int_{R_1+R_2} f(t, \phi, \phi_t)\, dt.$$

If we apply the mean value theorem to each of the three integrals and let b tend to zero, we find by (2) and (3)

$$\epsilon f(P) + (1-\epsilon) f\left(p - \frac{\epsilon}{1-\epsilon}\Delta p\right) - f(p) \geqq 0, \tag{4}$$

where for convenience the arguments $\bar{t}$ and $\bar{x} = \phi(\bar{t})$ of the function f have been omitted. In deriving the necessary condition (4) no use has been made of the existence of the partial derivatives f_i^α. When f has a total differential with respect to the arguments p_α^i, we may divide (4) by ϵ and let ϵ approach zero to obtain

$$f(P) - f(p) - (P_\alpha^i - p_\alpha^i) f_i^\alpha(p) \geqq 0. \tag{5}$$

2. **Transformation of the Weierstrass condition for the parametric problem.** When the variables t_α are regarded as parameters and only the x^i are considered as coördinates, we usually wish the integral I to be independent of parameter transformations with positive Jacobian determinant. A necessary and sufficient condition for this is that the integrand $f(t, x, p)$ be independent of the t_α, and be equal to a function $g(x, j)$ of the coördinates x^i and the n-rowed minors j_ρ of the matrix $p = (p_\alpha^i)$ which is positively homogeneous of the first degree in j. We assume that $n < m$ and that on admissible manifolds the matrix p has rank n at every point, so that not all of the j_ρ are zero. If, for each α, $(p_\alpha^1, \cdots, p_\alpha^m)$ are regarded as homogeneous coördinates of a hyperplane in $(m-1)$-dimensional projective space, then the j_ρ may be regarded as the homogeneous coördinates of an $(m-n-1)$-dimensional linear manifold in this space. For a full discussion from the geometric point of view of the relations holding between the j_ρ, the reader may refer to Bertini's *Geometria proiettiva degli iperspazi*, pp. 33–39. The discussion given below is in a form better suited to our purpose, and is perhaps briefer.

In each matrix such as (p_α^i) and (k_i^τ) we shall understand for definiteness that the subscript is the row-index and the superscript is the column-index. We recall that all the solutions of the equations

$$p_\alpha^i w_i = 0 \tag{6}$$

may be expressed in terms of the n-rowed minors j_ρ of the matrix p. The index ρ may be regarded as a symbol for a combination of n elements chosen from the set $\mu = (1, \cdots, m)$, and the index τ as a symbol for a combination of $n + 1$ such elements. Each column of the matrix $k = (k_i^\tau)$ is to be a solution of equations (6) having $(m - n - 1)$ zero elements, the remaining elements being minors j_ρ formed from the $(n + 1)$ columns of p corresponding to the elements of the combination τ, with proper signs attached. The matrix k is to include all such columns of solutions, so that it has $m!/[(m - n + 1)!(n + 1)!]$ columns. A definite law may be set down determining the order of the columns of p occurring in each minor j_ρ and determining the signs chosen for the j_ρ in each column of k. We suppose such a law chosen, so that the vector j and the matrix k are uniquely determined when p is given. Under a transformation of the parameters t_α whose Jacobian matrix is denoted by A, the matrix p is replaced by $P = Ap$, and the new matrix K corresponding to P is obtained from k by multiplying the elements of the latter by the determinant of A. Thus in particular p is not completely determined by k.

The rank of the matrix k is always exactly $(m - n)$. For it cannot be more, since the rank of p is n and the columns of k are solutions of equations (6). And it cannot be less, since every solution of (6) is expressible in terms of the columns of k. This statement about the rank of the matrix k yields the quadratic relationships which hold between the minors j_ρ.

We next wish to show that *if a vector $J = (J_\rho)$ is given such that the corresponding matrix $K = (K_i^\tau)$ has rank $(m - n)$, then there exists a corresponding matrix $P = (P_\alpha^i)$ of which the J_ρ are the n-rowed minors.* It is plain that there exists a matrix p of rank n such that $p_\alpha^i K_i^\tau = 0$. For definiteness of notation suppose the matrix K arranged in blocks in the form

$$\begin{pmatrix} K_1 & K_3 \\ K_2 & K_4 \end{pmatrix},$$

where K_1 is a square non-singular minor of $m - n$ rows, having zeros off the main diagonal, and diagonal elements equal to $\pm J_{\rho_0}$. Then p has the form (p_1, p_2), where p_2 has n columns and is non-singular. The block k_1 of the matrix

$$k = \begin{pmatrix} k_1 & k_3 \\ k_2 & k_4 \end{pmatrix}$$

formed from the matrix p is obviously a multiple cK_1 of K_1. Every column of k is a linear combination of the columns of

$$\begin{pmatrix} K_1 \\ K_2 \end{pmatrix}, \tag{7}$$

so that $k_2 = cK_2$. We next proceed to show that the elements of K_3 and hence of K_4 are determined by the elements of K_1 and K_2. Recall that K_1 has elements $\pm J_{\rho_0} \neq 0$ on the diagonal and zeros elsewhere. The row index i for K_1 ranges over the set $\mu - \rho_0$. The elements K_i^τ of K_2 are all the J_σ such that σ has $n - 1$ elements in common with ρ_0. If now τ contains just two elements of $\mu - \rho_0$ and i is in $\mu - \rho_0$, K_i^τ is either zero or $\pm J_\sigma$, where σ contains $n - 1$ elements of ρ_0, and so K_i^τ is already determined. The remaining elements of these columns are then also determined since every column is a linear combination of the columns of (7). The elements so far determined include all the J_σ such that σ has at least $n - 2$ elements in common with ρ_0. Proceeding thus, we see that all the elements of K are determined by those of its minor (7). Hence $k = cK$, and the desired matrix P may be obtained from p by dividing a row of the latter by c. Obviously P is determined up to a transformation of the form $P' = AP$, where A has determinant unity.

We note next that if the matrix $P - p$ has rank one, a transformation A of determinant unity may be applied so that $P = p$ except in one row. Then $n - 1$ rows of p are orthogonal to the columns of both k and K, where k corresponds to p and K to P, so that the rank of the matrix (k, K) is not greater than $m - n + 1$. On the other hand, suppose the matrices k and K are given, each of rank $m - n$, and that (k, K) has rank at most $m - n + 1$. Then we may select the matrices p and P corresponding to k and K, respectively, so that they differ at most in the first row, and then obviously $P - p$ has rank one at most.

In terms of matrices p and P so selected the Weierstrass condition of §1 yields the inequality

$$f(P) - f(p) - (P_1^i - p_1^i)f_i^1(p) \geqq 0.$$

When this is expressed in terms of the function $g(x, j)$, it becomes

$$g(J) - g(j) - (P_1^i - p_1^i)g_\rho \frac{\partial j_\rho}{\partial p_1^i} \geqq 0,$$

where $g_\rho = \partial g/\partial j_\rho$. .But $j_\rho = p_1^i \partial j_\rho/\partial p_1^i$, $J_\rho = P_1^i \partial j_\rho/\partial p_1^i$. Thus for parametric problems the Weierstrass condition may be stated in the following form:

$$E(x, j, J) = g(x, J) - g(x, j) - (J_\rho - j_\rho)g_\rho(x, j) \geqq 0$$

for every element (x, j) of the minimizing manifold and for every vector J such that the corresponding matrix K has rank $m - n$ and (k, K) has rank not greater than $m - n + 1$.

THE UNIVERSITY OF CHICAGO.

COMPLETELY MONOTONE FUNCTIONS AND SEQUENCES

By Willy Feller

1. **Introduction.** Consider a function $f(x)$ defined by the Laplace-Stieltjes integral

$$f(x) = \int_0^\infty e^{-xt}\, dF(t), \tag{1}$$

where $F(t)$ is never decreasing and is bounded in every finite interval and the integral converges, say, for $x > 0$. In this paper we propose to deduce a simple inversion formula for (1) by an argument which enables us at the same time to prove in a very natural way some theorems of the theory of Laplace-Stieltjes integrals. Thus our argument provides an extremely simple proof of the well-known theorem to the effect that a function $f(x)$ can be represented in the form (1) if, and only if, it is completely monotonic, i.e., if it has for $x > 0$ derivatives of any finite order such that

$$(-1)^n f^{(n)}(x) \geqq 0 \qquad (n = 0, 1, \cdots). \tag{2}$$

A theorem which is substantially equivalent to this statement is proved by F. Hausdorff.[1] In its present form it was first formulated by S. Bernstein,[2] and subsequently independently by D. V. Widder.[3] A simplified proof was then given by J. D. Tamarkin,[4] and subsequently Widder himself proposed an alternative proof.[5] Widder has also proposed some inversion formulas[6] for integrals of type (1), but the formula given in the sequel seems nevertheless to be of some interest and proves the theorem of Hausdorff-Bernstein in a more direct way. Moreover, the method is easily applicable to other problems, in particular to the problems treated by Widder in his papers referred to above. As an example we deduce a necessary and sufficient condition that a function $f(x)$ be representable in the form (1), when $F(t)$ is only supposed to be of bounded variation in every finite interval; this condition is equivalent to a similar condition obtained by Widder.[7] Also the case of a function $F(t)$ with a bounded derivative will be treated in the sequel.

We then deal with the general interpolation problem for completely mono-

Received April 1, 1939.

[1] Hausdorff [4], II, Theorem 3. (Numbers in brackets refer to the list of references at the end of the paper.)

[2] Bernstein [1].

[3] Widder [10].

[4] Tamarkin [9].

[5] Widder [11], Theorems 17–18.

[6] Widder [11].

[7] Widder [10], Theorem 12.

tonic functions. Let $\{x_n\}$ $(n = 0, 1, \cdots)$ denote an increasing sequence of non-negative numbers tending to infinity, and let to every x_n correspond a number $a_n \geqq 0$. It is moreover supposed that the series $\sum 1/x_n$ is divergent. Then a necessary and sufficient condition that there is a function $f(x)$, completely monotonic for $x > x_0$ and such that $f(x_n) = a_n$ $(n = 1, 2, \cdots)$ is that the divided differences of the sequence $\{a_n\}$, as defined by (24), be non-negative. In that case $f(x)$ is shown to be uniquely determined and representable by Newton's interpolation series. For $n = 0$ we always have $f(x_0) \leqq a_0$, but in general equality does not hold. The argument in this section will rest on quite elementary properties of the interpolation series. We notice that this interpolation problem has already been treated by Hausdorff[8] in connection with his investigations of the moment problem for $0 \leqq x \leqq 1$. Subsequently Widder[9] has by another method treated in detail the special case of equal intervals, $x_n = n$.

Finally, we derive an inversion formula which determines the function $F(t)$ of (1) directly in terms of an arbitrary sequence of values $f(x_n)$, provided only that x_n tends to infinity and the series $\sum 1/x_n$ diverges. This formula is a counterpart of the inversion formula mentioned above, the only difference being that Taylor's formula is replaced by Newton's interpolation series. In the special case of equal intervals, $x_n = n$, related inversion formulas have been found by Hausdorff and Widder,[10] but our formula uses other differences and reduces to neither of theirs.

Incidentally, it may be pointed out that the general inversion formula for unequal intervals of the argument is also of great importance for the stochastical theory of telephone traffic. In fact, this problem was proposed to the author by Conny Palm, of the telephone-administration of Stockholm, who was confronted with this and similar problems in the course of his important practical investigations on the telephone traffic. It seems that our inversion formulas are, with a slight modification, very suitable for numerical computations.[11]

2. **Theorem of Bernstein-Widder and inversion formula.** We proceed to prove

THEOREM 1. *A necessary and sufficient condition that the function $f(x)$ can be represented, for $x > 0$, in the form* (1), *with $F(t)$ non-decreasing and bounded in every finite interval, is that $f(x)$ have for $x > 0$ finite derivatives of all orders satisfying relation* (2). *In that case*

[8] Hausdorff [4], II.

[9] Widder [10], pp. 880–886, and [11], part V.

[10] Hausdorff [5] and Widder [11], pp. 174–194.

[11] Dr. O. Lundberg has kindly drawn my attention to the fact that a proof of the theorem of Hausdorff-Bernstein on lines similar to those indicated in the present paper was recently sketched by Dubourdieu [3]. His starting point is the differential equations of a particular stochastic process related to the theory of sickness-insurance, and he arrives in the particular case of a bounded $F(t)$ at the inversion formula (3) of our text and the proposition referred to. For an alternative proof in this particular case cf. also footnote 13.

$$F(t) = \lim_{\eta \to \infty} \sum_{n=0}^{[t\eta]} \frac{(-\eta)^n}{n!} f^{(n)}(\eta) \tag{3}$$

in any continuity point of $F(t)$.

The first part of the theorem, as stated above, is proved by Hausdorff, S. Bernstein and D. V. Widder. We note that the restriction to the interval $x > 0$ means no loss of generality. For if the integral (1) converges for $x > x_0 > 0$ only, we get for $f_1(x) = f(x + x_0)$ and $x > 0$ the representation

$$f_1(x) = \int_0^\infty e^{-xt}\, dF_1(t), \quad \text{with} \quad F_1(t) = \int_0^t e^{-x_0\tau}\, dF(\tau), \tag{4}$$

and conversely.

Proof of Theorem 1. The necessity of the condition (2) needs no comment. In order to prove the sufficiency we define a function $F_\eta(t)$, depending on a parameter $\eta > 0$, by $F_\eta(0) = 0$ and

$$F_\eta(t) = \sum_{n=0}^{[t\eta]} \frac{(-\eta)^n}{n!} f^{(n)}(\eta) \qquad \text{for } t > 0. \tag{5}$$

By (2) $F_\eta(t)$ is for $t \geqq 0$ a never decreasing function. With this function we have, since $f(x)$ is analytic for $x > 0$,[12]

$$\begin{aligned}\int_0^\infty e^{-xt}\, dF_\eta(t) &= \sum_{n=0}^{\infty} e^{-xn/\eta} \frac{(-\eta)^n}{n!} f^{(n)}(\eta) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} (\eta\{1 - e^{-x/\eta}\} - \eta)^n f^{(n)}(\eta) \\ &= f(\eta\{1 - e^{-x/\eta}\}).\end{aligned}$$

Thus

$$\lim_{\eta \to \infty} \int_0^\infty e^{-xt}\, dF_\eta(t) = f(x). \tag{6}$$

We shall now investigate the behavior of $F_\eta(t)$ for $\eta \to \infty$. If $f(0)$ is finite, it follows from (2) and (5) that $F_\eta(t) \leqq f(0)$ for all $\eta > 0$. Otherwise choose $\delta > 0$ arbitrarily small and denote by a a number such that $a > e^\delta$. Then for $0 \leqq n \leqq t\eta$ and η sufficiently large

$$\frac{(\eta - \delta)^n}{\eta^n} \geqq \left(1 - \frac{\delta}{\eta}\right)^{t\eta} \geqq a^{-t},$$

and hence

$$F_\eta(t) \leqq a^t \sum_{n=0}^{[t\eta]} \frac{(\eta - \delta)^n}{n!} (-1)^n f^{(n)}(\eta) \leqq a^t f(\delta). \tag{7}$$

It follows that it is possible to pick out a sequence $\eta_k \to \infty$ such that $F_{\eta_k}(t)$ con-

[12] See, for example, S. Bernstein, *Leçons sur les propriétés extrémales des fonctions analytiques d'une variable réelle*, Paris, 1926, p. 190.

verges essentially to a never decreasing function $F(t)$, which is bounded in every finite interval:

$$(8) \qquad \lim_{k\to\infty} F_{\eta_k}(t) = F(t)$$

in any continuity point of $F(t)$.

Next we prove that

$$(9) \qquad \lim_{k\to\infty} \int_0^\infty e^{-xt}\, dF_{\eta_k}(t) = \int_0^\infty e^{-xt}\, dF(t)$$

for $x > 0$. Now for any $T > 0$ we have

$$\int_{T+1}^\infty e^{-xt}\, dF_\eta(t) \leqq \sum_{n \geqq T_\eta} e^{-xn/\eta} \frac{(-\eta)^n}{n!} f^{(n)}(\eta) \leqq e^{-\frac{1}{2}xT} \sum_{n\geqq T_\eta} \frac{(-\eta e^{-\frac{1}{2}x/\eta})^n}{n!} f^{(n)}(\eta),$$

and hence by the argument already used

$$\int_{T+1}^\infty e^{-xt}\, dF_\eta(t) \leqq e^{-\frac{1}{2}xT} \sum_{n=0}^\infty \frac{(-\eta e^{-\frac{1}{2}x/\eta})^n}{n!} f^{(n)}(\eta)$$
$$= e^{-\frac{1}{2}xT} f(\eta\{1 - e^{-\frac{1}{2}x/\eta}\}) \leqq e^{-\frac{1}{2}xT} f(1 - e^{-\frac{1}{2}x}),$$

provided that $\eta > 1$. Consequently, given any $\epsilon > 0$ and $\delta > 0$, we can choose a $T > 0$ such that

$$\int_T^\infty e^{-xt}\, dF_\eta(t) < \epsilon$$

for $x > \delta$ and $\eta > 1$. Thus the integrals in (9) converge uniformly, so that (9) follows from (8).

Comparing (9) with (6), we get for $f(x)$ a representation of the desired type, and in order to complete the proof, it remains only to show that (8) holds for *any* sequence η_k tending with k to infinity. But this is almost obvious, for otherwise we should get two essentially different representations for $f(x)$ in the form (1), and this is known to be impossible.[13]

[13] This is a simple consequence of the theorem of Weierstrass on approximation by polynomials; see, e.g., Hausdorff [5], pp. 222-223. At the present stage of our proof the correctness of the statement may also be seen directly in the following way. It is known that

$$(*) \qquad \lim_{k\to\infty} e^{-kt} \sum_{r=0}^{k} \frac{(kt)^r}{r!} = \begin{cases} 1 & \text{if } t < 1, \\ 0 & \text{if } t > 1. \end{cases}$$

This is Theorem 1 of Widder [11]. It may also easily be verified by evaluating $e^{-kt}(kt)^k/k!$ by Stirling's formula and majorating $\sum_0^k (kt)^r/r!$ if $t > 1$ or $\sum_{k+1}^\infty (kt)^r/r!$ if $t < 1$ by a geometrical series. From (*) it follows that

$$\lim_{x\to\infty} e^{-xt} \sum_{[t_1 x]+1}^{[t_2 x]} \frac{(xt)^r}{r!} = 1 \quad \text{if} \quad t_1 < t < t_2.$$

3. The function $F(t)$ of bounded variation.

THEOREM 2. *A necessary and sufficient condition that $f(x)$ can, for $x > 0$, be represented in the form*

$$f(x) = \int_0^\infty e^{-xt}\,dV(t), \tag{10}$$

with $V(t)$ of bounded variation in every finite interval and the integral converging absolutely for $x > 0$, is that for any $\delta > 0$ there exists a constant $M = M(\delta)$ such that

$$\sum_{n=0}^\infty \frac{(x-\delta)^n}{n!}\,|f^{(n)}(x)| < M(\delta) \tag{11}$$

uniformly for $x > \delta$.

This condition is easily shown to be equivalent to a similar condition derived by Widder.[14]

Proof. (i) Let $f(x)$ be defined by (10) and suppose that

$$V(t) = V_1(t) - V_2(t) \tag{12}$$

is the canonical decomposition of $V(t)$ in two non-decreasing functions. Define $f_1(x)$ and $f_2(x)$ by

$$f_i(x) = \int_0^\infty e^{-xt}\,dV_i(t) \qquad (i = 1, 2) \tag{13}$$

and put

$$\varphi(x) = \int_0^\infty e^{-xt}\,d\{V_1(t) + V_2(t)\}. \tag{14}$$

Then, for any $x > \delta > 0$ and any $N > 0$,

$$\begin{aligned} 0 \leqq \sum_{n=0}^N \left| \frac{(x-\delta)^n}{n!}\,\varphi^{(n)}(x) \right| &= \int_0^\infty e^{-xt} \sum_{n=0}^N \frac{(x-\delta)^n}{n!}\,t^n\,d\{V_1(t) + V_2(t)\} \\ &\leqq \int_0^\infty e^{-\delta t}\,d\{V_1(t) + V_2(t)\} = \varphi(\delta). \end{aligned} \tag{15}$$

Now let $f(x) = \int_0^\infty e^{-xt}dG(t)$, with $G(t)$ never decreasing and bounded in every finite interval. Then

$$\lim_{x\to\infty} \sum_{[t_1x]+1}^{[t_2x]} \frac{(-x)^r f^{(r)}(x)}{r!} \geqq \lim_{x\to\infty} \int_{t_1}^{t_2} e^{-xt} \sum_{[t_1x]+1}^{[t_2x]} \frac{(xt)^r}{r!}\,dG(t) = G(t_2) - G(t_1)$$

for any pair (t_1, t_2) of continuity points of $G(t)$. Hence $G(t_2) - G(t_1) \leqq F(t_2) - F(t_1)$, where $F(t)$ is defined by (8), and this implies the uniqueness of the representation (1). In case of a bounded $F(t)$ the argument yields another simple proof of the inversion formula (3) and the theorem of Hausdorff-Bernstein.

[14] Widder [10], Theorem 12.

But since

$$\varphi^{(n)}(x) = f_1^{(n)}(x) + f_2^{(n)}(x) \quad \text{and} \quad f_1^{(n)}(x) f_2^{(n)}(x) \geqq 0,$$

obviously

$$| f^{(n)}(x) | = | f_1^{(n)}(x) - f_2^{(n)}(x) | \leqq | \varphi^{(n)}(x) | .$$

Thus (15) implies (11) with $M = \varphi(\delta)$.

(ii) Conversely, let us suppose that (11) holds. Define two non-decreasing functions $P_\eta(t)$ and $N_\eta(t)$, depending on a parameter η, by

$$\begin{aligned} P_\eta(t) &= \sum_{n=0}^{[t\eta]} \frac{\eta^n}{n!} | f^{(n)}(\eta) |, \\ N_\eta(t) &= \sum_{n=0}^{[t\eta]} \frac{\eta^n}{n!} \{ | f^{(n)}(\eta) | - (-1)^n f^{(n)}(\eta) \}, \end{aligned} \tag{16}$$

where $t > 0$. For $P_\eta(t)$ we may by an argument already used deduce an inequality similar to (7), viz.,

$$P_\eta(t) \leqq a^t M(\delta), \tag{17}$$

holding for any given $\delta > 0$, a satisfying the inequality $a > e^\delta$, and η sufficiently large. Hence for a suitable sequence η_k tending to infinity the functions $P_{\eta_k}(t)$ converge essentially to a never decreasing function $P(t)$ which is bounded in any finite interval. Now by (16) the variation of $N_\eta(t)$ in any finite interval does not exceed twice the corresponding variation of $P_\eta(t)$. Hence we may choose the sequence η_k in a way that also $N_{\eta_k}(t)$ converges essentially to a non-decreasing function $N(t)$.

Repeating then the argument used for the proof of Theorem 1, we see that for $\eta > 1$

$$\begin{aligned} \int_{T+1}^{\infty} e^{-xt} \, dP_\eta(t) &\leqq \sum_{n \geqq T\eta} e^{-xn/\eta} \frac{\eta^n}{n!} | f^{(n)}(\eta) | \leqq e^{-\frac{1}{2}xT} \sum_{n \geqq T\eta} \frac{(\eta e^{-\frac{1}{2}x/\eta})^n}{n!} | f^{(n)}(\eta) | \\ &\leqq e^{-\frac{1}{2}xT} \sum_{n=0}^{\infty} \frac{| \eta\{1 - e^{-\frac{1}{2}x/\eta}\} - \eta |^n}{n!} | f^{(n)}(\eta) | \leqq e^{-\frac{1}{2}xT} M(\eta\{1 - e^{-\frac{1}{2}x/\eta}\}) \\ &\leqq e^{-\frac{1}{2}xT} M(1 - e^{-\frac{1}{2}x}). \end{aligned}$$

Hence it is seen that the integrals converge uniformly and

$$\lim_{k \to \infty} \int_0^\infty e^{-xt} \, dP_{\eta_k}(t) = \int_0^\infty e^{-xt} \, dP(t),$$

and also, since the variation of $N_\eta(t)$ is dominated by that of $2P_\eta(t)$,

$$\lim_{k \to \infty} \int_0^\infty e^{-xt} \, dN_{\eta_k}(t) = \int_0^\infty e^{-xt} \, dN(t).$$

We thus obtain

$$\int_0^\infty e^{-xt} d\{P(t) - N(t)\} = \lim_{k\to\infty} \int_0^\infty e^{-xt} d\{P_{\eta_k}(t) - N_{\eta_k}(t)\}$$

$$= \lim_{k\to\infty} \sum_{n=0}^{\infty} e^{-xn/\eta_k} \frac{(-\eta_k)^n}{n!} f^{(n)}(\eta_k) = \lim_{k\to\infty} \sum_{n=0}^{\infty} \frac{1}{n!} (\eta_k\{1 - e^{-x/\eta_k}\} - \eta_k)^n f^{(n)}(\eta_k).$$

But $f(x)$ is clearly an analytic function[15] so that the last series equals

$$f(\eta_k\{1 - e^{-x/\eta_k}\})$$

which quantity tends to $f(x)$. Hence

$$f(x) = \int_0^\infty e^{-xt} d\{P(t) - N(t)\}, \tag{18}$$

and the proof is complete.

4. **The function $F(t)$ with a bounded derivative.** As a further application of the method we prove the following theorem of Widder:[16]

THEOREM 3. *A necessary and sufficient condition that $f(x)$ can be expressed in the form*

$$f(x) = \int_0^\infty e^{-xt} \varphi(t)\, dt, \tag{19}$$

with $\varphi(t)$ uniformly bounded, is that

$$\left| \frac{x^n}{n!} f^{(n)}(x) \right| < \frac{A}{x} \qquad (n = 0, 1, \cdots), \tag{20}$$

where A is a constant.

Proof. (i) If $f(x)$ is given by (19) and $|\varphi(x)| < A$, then

$$\left| \frac{x^n}{n!} f^{(n)}(x) \right| = \left| \int_0^\infty e^{-xt} \frac{(xt)^n}{n!} \varphi(t)\, dt \right| \leqq A \int_0^\infty e^{-xt} \frac{(xt)^n}{n!}\, dt = \frac{A}{x}.$$

This proves the necessity of the condition.

(ii) Conversely, it follows from (20) that for $0 < t_1 < t_2$

$$\sum_{[t_1\eta]}^{[t_2\eta]} \frac{\eta^n}{n!} \left| f^{(n)}(\eta) \right| < A(t_2 - t_1). \tag{21}$$

Consider now the function $P_\eta(t)$ defined by (16). By (21) its variation in any interval $0 < t_1 < t_2$ does not exceed $A(t_2 - t_1)$, and the same is naturally true also for the limiting function $P(t)$. Thus the derivatives of $P(t)$ are bounded by A, and hence

$$P(t) = \int_0^t \varphi_1(s)\, ds \quad \text{with} \quad 0 \leqq \varphi_1(s) \leqq A. \tag{22}$$

[15] See footnote 12.

[16] Widder [10], Theorem 13.

But, as it has already been observed, the variation of $N_\eta(t)$ in any interval does not exceed twice the corresponding variation of $P_\eta(t)$, so that also for $N(t)$ a relation similar to (22) holds. In view of (18) these relations imply (19).

5. **The interpolation problem.** Consider a sequence of real numbers

$$(23)\qquad 0 \leqq x_0 < x_1 < \cdots < x_n < \cdots, \qquad x_n \to \infty$$

and let us suppose that to each x_n corresponds a number $a_n \geqq 0$. Then the *divided differences* of the sequence $\{a_n\}$ are formed by induction in the following manner:[17]

$$(24)\qquad \begin{gathered} [a_n] = a_n\,, \\ [a_{i_0}, a_{i_1}, \cdots, a_{i_n}] = \frac{[a_{i_1}, a_{i_2}, \cdots, a_{i_n}] - [a_{i_0}, a_{i_1}, \cdots, a_{i_{n-1}}]}{x_{i_n} - x_{i_0}}, \end{gathered}$$

so that the order of a divided difference is less by unity than the number of arguments required for its definition. It should be remembered that the divided differences depend essentially on both sequences $\{x_n\}$ and $\{a_n\}$ even though the usual notation $[a_{i_0}, \cdots, a_{i_n}]$ does not show it. We note further that the divided differences are symmetrical functions of their arguments. In case of equal intervals of the argument, say $x_n = n$, we have

$$(25)\qquad [a_0, a_1, \cdots, a_n] = \frac{1}{n!}\Delta^n a_0 = \frac{1}{n!}\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}a_k\,.$$

For the sequel it will, however, be necessary not to restrict ourselves to the divided differences $[a_k, a_{k+1}, \cdots, a_n]$ of consecutive a_i.

The sequence $\{a_n\}$, depending on the argument-values x_n, will be called *completely monotonic*, if for any set of non-negative integers $(i_0, \cdots, i_n)$ $(n = 0, 1, \cdots)$

$$(26)\qquad (-1)^n[a_{i_0}, a_{i_1}, \cdots, a_{i_n}] \geqq 0.$$

Consider now a sequence $\{a_n\}$ defined by the values taken by a completely monotonic function at the points x_n:

$$(27)\quad a_n = f(x_n) \qquad (n = 0, 1, \cdots); \qquad (-1)^n f^{(n)}(x) \geqq 0 \quad \text{for} \quad x \geqq x_0\,.$$

Then it is known that there is some value ξ between x_{i_0} and x_{i_n} such that

$$[a_{i_0}, \cdots, a_{i_n}] = \frac{1}{n!}f^{(n)}(\xi),$$

and hence the sequence $\{a_n\}$ is completely monotonic. Before proving the converse theorem, let us observe that increasing a_0 makes of $\{a_n\}$ a new sequence which is still completely monotonic. On the other hand, it will be proved that decreasing a_0 in a sequence such as (27) makes of it a sequence which is no longer

[17] Cf., e.g., the treatise of Milne-Thomson [7].

completely monotonic. Such sequences are called by Widder *minimal sequences.*[18] With this preliminary remark we proceed to prove

THEOREM 4. *If the sequence* $\{a_n\}$ *corresponding to the argument values* (23) *is completely monotonic, and if the series*

$$\sum 1/x_n \tag{28}$$

diverges, then there is a uniquely determined function $f(x)$, *defined and completely monotonic for* $x \geqq x_0$, *such that*

$$f(x_n) = a_n \qquad (n = 1, 2, \cdots); \tag{29}$$

for this function

$$f(x_0) \leqq a_0, \tag{30}$$

and if $i_0, i_1, \cdots$ *is any strictly increasing sequence of integers such that* $\sum 1/x_{i_n}$ *diverges, then for* $x \geqq x_0$

$$f(x) = \sum_{k=0}^{\infty} (-1)^k [a_{i_0}, a_{i_1}, \cdots, a_{i_k}] (x_{i_0} - x) \cdots (x_{i_{k-1}} - x). \tag{31}$$

Proof. Let $\{i_n\}$ be a strictly increasing sequence of positive integers and let

$$\sum 1/x_{i_n} \tag{32}$$

be divergent. Suppose that $i_0 \geqq 1$, and let i be an integer such that $0 \leqq i < i_0$. Then by Newton's interpolation formula (or by a straightforward inductive argument)

$$\begin{aligned} &\sum_{k=0}^{N} (-1)^k [a_{i_0}, a_{i_1}, \cdots, a_{i_k}] (x_{i_0} - x_i) \cdots (x_{i_{k-1}} - x_i) \\ &\qquad = a_i - (-1)^{N+1} [a_i, a_{i_0}, a_{i_1}, \cdots, a_{i_N}] (x_{i_0} - x_i) \cdots (x_{i_N} - x_i) \end{aligned} \tag{33}$$

for any $N \geqq 0$. Now by (26) the sum on the left side contains non-negative terms only, while the right member does not exceed a_i. Hence we may let $N \to \infty$ and obtain

$$\sum_{k=0}^{\infty} (-1)^k [a_{i_0}, a_{i_1}, \cdots, a_{i_k}] (x_{i_0} - x_i) \cdots (x_{i_{k-1}} - x_i) \leqq a_i. \tag{34}$$

But for any fixed $x > x_i$ and k sufficiently large $0 < x_{i_k} - x < x_{i_k} - x_i$, so that (34) implies that the series

$$\sum_{k=0}^{\infty} (-1)^k [a_{i_0}, a_{i_1}, \cdots, a_{i_k}] (x_{i_0} - x) \cdots (x_{i_{k-1}} - x) \tag{35}$$

converges for all $x \geqq x_0$.

[18] Widder [10], p. 880.

Next we show that in case $1 \leqq i < i_0$

$$(36)\qquad (-1)^{N+1}[a_i, a_{i_0}, a_{i_1}, \cdots, a_{i_N}](x_{i_0} - x_i) \cdots (x_{i_N} - x_i) \to 0$$

as $N \to \infty$. In fact, by (33) and (26) the quantity on the left side is never increasing and non-negative. Suppose now that

$$(37)\qquad \lim_{N\to\infty} (-1)^{N+1}[a_i, a_{i_0}, \cdots, a_{i_N}]\,(x_{i_0} - x_i) \cdots (x_{i_N} - x_i) = a > 0.$$

Then, for any $k \geqq 0$, we should have

$$(38)\qquad \begin{aligned} &(-1)^{k+1}[a_i, a_{i_0}, a_{i_1}, \cdots, a_{i_k}]\,(x_i - x_0)(x_{i_0} - x_0) \cdots (x_{i_{k-1}} - x_0) \\ &\geqq (-1)^{k+1}[a_i, a_{i_0}, \cdots, a_{i_k}]\,(x_{i_0} - x_i) \cdots (x_{i_{k-1}} - x_i)(x_{i_k} - x_i)\,\frac{x_i - x_0}{x_{i_k} - x_i} \\ &\geqq a\,\frac{x_i - x_0}{x_{i_k} - x_i}. \end{aligned}$$

But it was supposed that $i \geqq 1$; hence it follows from (34) by a slight change of notation that

$$\sum_{k=0}^{\infty} (-1)^{k+1}[a_i, a_{i_0}, a_{i_1}, \cdots, a_{i_k}]\,(x_i - x_0)(x_{i_0} - x_0) \cdots (x_{i_{k-1}} - x_0) \leqq f(x_0),$$

while

$$\sum_{k=0}^{\infty} \frac{a(x_i - x_0)}{x_{i_k} - x_i}$$

diverges by hypothesis (32). This clearly contradicts (38) so that (37) must be false. Thus (36) is proved.

Comparing now (36) with (33), we see that for $1 \leqq i < i_0$

$$(39)\qquad \sum_{k=0}^{\infty} (-1)^k [a_{i_0}, \cdots, a_{i_k}]\,(x_{i_0} - x_i) \cdots (x_{i_{k-1}} - x_i) = a_i .$$

For $i = 0$ the inequality (34) cannot, of course, be improved, as the left member is independent of a_0 while increasing a_0 does not alter the completely monotonic character of the sequence $\{a_n\}$.

Consider now the function defined by the series (35). We propose to show that it in no way depends upon the particular choice of the i_k. To prove this we put for $x \geqq x_0$

$$(40)\qquad f_r(x) = \sum_{k=0}^{\infty} (-1)^k [a_{i_r}, a_{i_{r+1}}, \cdots, a_{i_{r+k}}]\,(x_{i_r} - x) \cdots (x_{i_{r+k-1}} - x).$$

It is easy to prove that all $f_r(x)$ are identical. In fact, for any $r \geqq 0$ and $N \geqq 0$ we have

$$
\begin{aligned}
&\sum_{k=0}^{N} (-1)^k [a_{i_{r+1}}, a_{i_{r+2}}, \cdots, a_{i_{r+k+1}}] (x_{i_{r+1}} - x) \cdots (x_{i_{r+k}} - x) \\
&\quad = \sum_{k=0}^{N} (-1)^k \{[a_{i_r}, \cdots, a_{i_{r+k+1}}] (x_{i_{r+k+1}} - x_{i_r}) \\
&\qquad\qquad + [a_{i_r}, \cdots, a_{i_{r+k}}]\} (x_{i_{r+1}} - x) \cdots (x_{i_{r+k}} - x) \\
&\quad = -\sum_{k=1}^{N+1} (-1)^k [a_{i_r}, \cdots, a_{i_{r+k}}] (x_{i_{r+1}} - x) \cdots (x_{i_{r+k-1}} - x)(x_{i_{r+k}} - x_{i_r}) \\
&\qquad\qquad + \sum_{k=1}^{N} (-1)^k [a_{i_r}, \cdots, a_{i_{r+k}}] (x_{i_{r+1}} - x) \cdots (x_{i_{r+k}} - x) \\
&\quad = \sum_{k=0}^{N} (-1)^k [a_{i_k}, \cdots, a_{i_{r+k}}] (x_{i_r} - x) \cdots (x_{i_{r+k-1}} - x) - R_N,
\end{aligned}
\tag{41}
$$

where

$$R_N = (-1)^{N+1} [a_{i_r}, \cdots, a_{i_{r+N+1}}] (x_{i_{r+1}} - x) \cdots (x_{i_{r+N}} - x)(x_{i_{r+N+1}} - x_{i_r}).$$

But $R_N \to 0$ for any fixed $x \geqq x_0$ and r; this follows immediately if in (36) we replace i by i_k and i_k by i_{r+k+1} $(k = 0, 1, \cdots)$, which is only a change of notation. Thus we may in (41) let $N \to \infty$, and it is seen that

$$f_{r+1}(x) = f_r(x) = f(x). \tag{42}$$

Now it follows from (39) that $f(x_i) = a_i$ for $i = 1, 2, \cdots, i_r - 1$. But $i_r \to \infty$ as $r \to \infty$ and thus we conclude from (42) that $f(x_i) = a_i$ for $i \geqq 1$. Finally, $f(x_0) \leqq a_0$ by (34).

In order to complete the proof it remains only to show that $f(x)$ is completely monotonic and uniquely determined. But by writing $f(x)$ in the form (40), we readily see that for $x_0 \leqq x \leqq x_{i_r}$ and any $n \geqq 0$

$$(-1)^n f^{(n)}(x) \geqq 0, \tag{43}$$

since all the factors $x_{i_s} - x$ are non-negative and by hypothesis (26) holds. Now r is arbitrary and $x_{i_r} \to \infty$, so that (43) holds for any $x \geqq x_0$.

To prove the uniqueness of our interpolatory function we observe that (39) implies that two completely monotone sequences which coincide at the points of a sequence x_{i_k} such that $\sum 1/x_{i_k}$ diverges differ at most by the first term. On the other hand, if $f(x)$ is completely monotonic, and if $\{\xi_k\}$ is an arbitrary sequence of points, then the sequence $f(\xi_k)$ is also completely monotonic. Thus two completely monotonic functions which coincide at the points (23) are identical. This completes the proof of the theorem and we have also the following

COROLLARY. *Two completely monotone functions which coincide at the points of a sequence x_n such that $\sum 1/x_n$ diverges are identical.*

By Theorem 2 this may be expressed also in the following way. If

$$f(x) = \int_0^\infty e^{-xt}\, dV(t),$$

where $V(t)$ is of bounded variation in every finite interval and the integral converges absolutely for $x > x_0$, and if x_n are the zeros of $f(x)$, then the series $\sum 1/x_n$ converges. This theorem has been proved by Wintner,[19] who has pointed out that it is a consequence of a theorem of Müntz[20] to the effect that an arbitrarily good approximation of any continuous function in $0 \leqq x \leqq 1$ may be performed by linear aggregates of the functions x^{x_k} ($x_k \geqq 0$) provided that $\sum 1/x_k$ is divergent. Conversely, the possibility of such an approximation is a consequence of the uniqueness theorem, as has been proved here by a direct interpolation.

6. **Inversion formula for sequences.** Given a completely monotonic sequence $\{a_n\}$, by (31) we are able to calculate the corresponding interpolatory function $f(x)$, and then by (3) also the function $F(t)$ to which it belongs. It has, however, been pointed out in the introduction that practical problems require a direct determination of $F(t)$ in terms of a_n. For the special case of equal intervals, $x_n = n$, such inversion formulas have been derived by Hausdorff and Widder, but they are insufficient for our actual purposes.

Without loss of generality we may assume that $x_0 = 0$. We then have

THEOREM 5. *Let $0 = x_0 < x_1 < \cdots < x_n \to \infty$ be a sequence of points such that $\sum 1/x_n$ diverges but $\sum 1/x_n^2$ converges.*[21] *Let a completely monotonic sequence of numbers $a_n \geqq 0$ correspond to the points x_n. Then the completely monotonic function $f(x)$, which at the points x_n ($n = 1, 2, \cdots$) takes on the values a_n, may be represented in the form* (1) *with*

$$(44)\qquad F(t) = \lim_{N\to\infty} \sum_{k=0}^{\varphi_N(t)} (-1)^k [a_N, a_{N+1}, \cdots, a_{N+k}]\, x_N x_{N+1} \cdots x_{N+k-1},$$

where $\varphi_N(t)$ denotes the integer determined by the relations

$$(45)\qquad \sum_{k=N}^{\varphi_N(t)+N-1} \frac{1}{x_k} < t, \qquad \sum_{k=N}^{\varphi_N(t)+N} \frac{1}{x_k} \geqq t, \qquad \varphi_N(0) = -1.$$

Proof. Define for any integer $N > 0$ and for $t \geqq 0$ a function $F_N(t)$ by

$$(46)\qquad F_N(t) = \sum_{k=0}^{\varphi_N(t)} (-1)^k [a_N, a_{N+1}, \cdots, a_{N+k}]\, x_N x_{N+1} \cdots x_{N+k-1}.$$

[19] Wintner [12].

[20] Müntz [8] or, e.g., the treatise of Kaczmarz-Steinhaus [6].

[21] The supposition that $\sum 1/x_n^2$ is convergent means obviously no restriction.

Since the sequence $\{a_n\}$ is supposed to be completely monotonic, it is seen by (26) that $F_N(t)$ is a never decreasing function of t. Taking in (34) $x_i = x_0 = 0$, we see further that for any $N > 0$

$$F_N(t) \leqq a_0 .$$

Hence it is possible to pick out a subsequence $\{N_r\}$ such that the sequence $\{F_{N_r}(t)\}$ converges essentially to a never decreasing function $F(t)$. Since the functions $F_N(t)$ are uniformly bounded, we have for $x \geqq 0$

$$\lim_{r\to\infty} \int_0^\infty e^{-xt}\, dF_{N_r}(t) = \int_0^\infty e^{-xt}\, dF(t). \tag{47}$$

Now

$$\int_0^\infty e^{-xt}\, dF_N(t)$$
$$= \sum_{k=0}^{\infty} \exp\left(-x \sum_{\nu=N}^{N+k-1} \frac{1}{x_\nu}\right)(-1)^k[a_N, a_{N+1}, \cdots, a_{N+k}]x_N x_{N+1} \cdots x_{N+k-1}.$$

But since $\sum 1/x_\nu^2$ converges we have

$$\exp\left(-x \sum_{\nu=N}^{N+k-1} \frac{1}{x_\nu}\right) = \prod_{\nu=N}^{N+k-1}\left(1 - \frac{x}{x_\nu}\right) + O\left(\sum_{\nu=N}^{\infty} \frac{1}{x_\nu^2}\right).$$

Thus we obtain by (31)

$$\int_0^\infty e^{-xt}\, dF_N(t) = \epsilon_N f(0)$$
$$+ \sum_{k=0}^{\infty} (-1)^k[a_N, a_{N+1}, \cdots, a_{N+k}](x_N - x) \cdots (x_{N+k-1} - x)$$
$$= \epsilon_N f(0) + f(x),$$

with ϵ_N tending to zero as $N \to \infty$. Hence by (47)

$$\int_0^\infty e^{-xt}\, dF(t) = \lim_{r\to\infty} \int_0^\infty e^{-xt}\, dF_{N_r}(t) = f(x),$$

and this is a representation of the desired type. Now the representation (1) was seen to be unique, and it is therefore impossible to pick out two subsequences $\{F_{N_r}(t)\}$ with different limiting functions. This proves the correctness of (44) and completes the proof of the theorem.

Remark. Obviously we are free to use instead of $\varphi_N(t)$ any integer-valued function which as $N \to \infty$ is asymptotically equivalent to it. Thus we may in the case of equal interval, $x_n = nh$, write (46) in the form

$$F_N(t) \sum_{k=0}^{[N(e^{ht}-1)]} (-1)^k \binom{N+k-1}{k} \Delta^k a_N. \tag{48}$$

Replacing here N by a new parameter ηh, we readily see that as $h \to 0$ $F_N(t)$ tends to the function defined by (3).

In order to compare (48) with the corresponding functions of Hausdorff and Widder we make the substitution $z = e^{-t}$, thus getting a non-decreasing function in $0 < z \leqq 1$. Moreover, we have to take $h = 1$. Then Hausdorff's N-th approximating function is a step function with N jumps of amount $(-1)^{N-k}\binom{N}{k}$ $\Delta^{N-k}a_k$ at the points k/N ($k = 0, 1, \cdots, N$). Widder's N-th approximating function has, as well as ours, infinitely many jumps at the points $N/(N + k)$ ($k = 0, 1, \cdots$). But the amounts of the corresponding jumps differ: with Widder they are $(-1)^N\binom{N+k-1}{N}\Delta^N a_k$, and thus differences of a fixed order are used. With our formula the jumps are of amount $(-1)^k\binom{N+k-1}{k}\Delta^k a_N$, so that differences belonging to a fixed point are used, i.e., the differences entering also in Newton's interpolation series.

References

1. S. Bernstein, *Sur les fonctions absolument monotones*, Acta Mathematica, vol. 52(1929), pp. 1–66.
2. G. Doetsch, *Theorie und Anwendung der Laplace-Transformation*, Berlin, 1937.
3. J. Dubourdieu, *Les fonctions absolument monotones et la théorie mathématique de l'assurance-accidents*, Comptes Rendus de l'Acad. des Sciences, Paris, vol. 206(1938), pp. 556–557.
4. F. Hausdorff, *Summationsmethoden und Momentenfolgen*, I–II, Mathematische Zeitschrift, vol. 9(1921), pp. 74–109, 280–299.
5. F. Hausdorff, *Momentenprobleme für ein endliches Intervall*, ibid., vol. 16(1923), pp. 220–248.
6. St. Kaczmarz and H. Steinhaus, *Theorie der Orthogonalreihen*, Warszawa-Lwów, 1935.
7. L. M. Milne-Thomson, *The Calculus of Finite Differences*, London, 1933.
8. Ch. H. Müntz, *Ueber den Approximationssatz von Weierstrass*, Math. Abhandlgen. H. A. Schwarz gewidmet, Berlin, 1914, pp. 303–312.
9. J. D. Tamarkin, *On a theorem of Bernstein-Widder*, Trans. Amer. Math. Soc., vol. 33(1931), pp. 893–896.
10. D. V. Widder, *Necessary and sufficient conditions for the representation of a function as a Laplace integral*, Trans. Amer. Math. Soc., vol. 33(1931), pp. 851–892.
11. D. V. Widder, *The inversion of the Laplace integral and the related moment problem*, ibid., vol. 36(1936), pp. 107–200.
12. A. Wintner, *Bemerkung zum Eindeutigkeitssatz der Laplaceschen Transformierten*, Math. Zeitschrift, vol. 36(1933), pp. 638–640.

Stockholm.

ISOPERIMETRIC PROBLEMS OF BOLZA IN NON-PARAMETRIC FORM

By William T. Reid

1. Introduction. For isoperimetric problems of the calculus of variations in parametric form an elegant proof of sufficient conditions for a strong relative minimum is afforded by the use of the so-called Lindeberg theorem. This result was obtained by Lindeberg [8][1] in 1909, who at that time applied it to the plane isoperimetric problem in parametric form.[2] Subsequently, this theorem was generalized and extended by Levi [7] and Tonelli ([13]; [14], vol. 1, p. 321). Recently Perlin [9] has developed a generalized Lindeberg theorem and applied it to the study of sufficient conditions for the parametric problem of Lagrange involving isoperimetric side conditions.

For non-parametric isoperimetric problems of the calculus of variations, however, neither the result obtained by Lindeberg [8], nor any one of the extended forms of the Lindeberg theorem established by Levi [7] and Tonelli ([13]; [14], vol. 1, p. 422), is effective in the proof of sufficient conditions for a strong relative minimum. Recently, Hestenes [5] has given a sufficiency proof for the isoperimetric problem of Bolza in non-parametric form. His proof is a generalization of the usual field method and does not use any analogue of the Lindeberg theorem; in particular, it involves the breaking up of the given extremal into suitable subarcs.

It is the purpose of the present paper to derive an effective Lindeberg theorem for non-parametric problems of the calculus of variations with isoperimetric side conditions. For the sake of generality, we consider specifically a problem of Bolza with variable end-points. It is of interest to note that the analogue of the Lindeberg theorem presented in §4 involves only the Weierstrass $\mathcal{E}$-function associated with the problem under discussion. It is first proved in §3 that if E is a non-singular extremal for the given problem satisfying the Weierstrass condition II_N, then there exists an associated problem which is entirely equivalent to the initial problem, and for which new and stronger forms of the usual Clebsch and Weierstrass conditions hold. The forms of these conditions thus derived render a simplicity to the results of §4 comparable to that for a problem which involves no auxiliary differential equations. They also enable one to simplify the expansion proof of sufficient conditions for the non-parametric problem of Bolza given by the author ([10], [11]).

Received April 24, 1939; presented to the American Mathematical Society, December 28, 1938.

[1] Numbers in brackets refer to the bibliography at the end of this paper.

[2] See also Bolza [2], pp. 515–518.

Finally, in §6 there is established an Osgood theorem for the calculus of variations problem herein considered.

2. Formulation of the problem. The problem to be considered is that of finding in a class of arcs

$$y_i(x) \qquad (i = 1, \cdots, n; x_1 \leqq x \leqq x_2) \tag{2.1}$$

satisfying the auxiliary conditions

$$\phi_\alpha[x, y, y'] = 0 \qquad (\alpha = 1, \cdots, m < n), \tag{2.2}$$

$$\psi_\mu[x_1, y(x_1), x_2, y(x_2)] = 0 \qquad (\mu = 1, \cdots, p \leqq 2n + 2), \tag{2.3}$$

$$\chi_s[x_1, y(x_1), x_2, y(x_2)] + \int_{x_1}^{x_2} h_s[x, y, y']\, dx = 0 \qquad (s = 1, \cdots, q) \tag{2.4}$$

one which minimizes a given functional

$$J = g[x_1, y(x_1), x_2, y(x_2)] + \int_{x_1}^{x_2} f[x, y, y']\, dx. \tag{2.5}$$

This is a problem of the general type of Bolza in the calculus of variations involving the isoperimetric conditions (2.4). For simplicity, this problem will be referred to as B.

It is supposed that there is given an open region R of points $[x, y, r]$ in which the functions $f[x, y, r]$, $\phi_\alpha[x, y, r]$, $h_s[x, y, r]$ are of class C^3. We shall also suppose that R_1 is an open region of $(2n + 2)$-dimensional sets $[x_1, y_{i1}, x_2, y_{i2}]$ in which the functions g, ψ_μ, χ_s are of class C^2 and, moreover, the matrix $(\psi_{\mu x_1}\ \psi_{\mu y_{i1}}\ \psi_{\mu x_2}\ \psi_{\mu y_{i2}})$ is of rank p. An arc (2.1) will be said to be *admissible* for problem B if its defining functions are continuous, possess piecewise continuous derivatives of the first order, its elements $[x, y(x), y'(x)]$ are in R, its end-points $[x_1, y_i(x_1), x_2, y_i(x_2)]$ lie in R_1, and for this arc conditions (2.2), (2.3), and (2.4) hold.

It is readily seen that the above problem B is formally equivalent to a problem of Bolza of the usual sort involving $n + q$ dependent functions. For if we set

$$u_s(x) = \int_x^{x_2} h_s[t, y(t), y'(t)]\, dt \qquad (s = 1, \cdots, q), \tag{2.6}$$

the above problem becomes that of finding in a class of arcs

$$y_i(x), \quad u_s(x) \qquad (i = 1, \cdots, n; s = 1, \cdots, q; x_1 \leqq x \leqq x_2) \tag{2.1'}$$

satisfying

$$\phi_\alpha[x, y, y'] = 0, \qquad u_s' + h_s[x, y, y'] = 0 \quad (\alpha = 1, \cdots, m; s = 1, \cdots, q), \tag{2.2'}$$

$$\begin{aligned} &\psi_\mu[x_1, y(x_1), x_2, y(x_2)] = 0 && (\mu = 1, \cdots, p), \\ &\chi_s[x_1, y(x_1), x_2, y(x_2)] + u_s(x_1) = 0, \quad u_s(x_2) = 0 && (s = 1, \cdots, q) \end{aligned} \tag{2.3'}$$

one which minimizes the functional (2.5). This latter problem will be referred to as $\mathfrak{B}$. An arc (2.1′) will be said to be admissible for $\mathfrak{B}$ if and only if it is related by equations (2.6) to an admissible arc for B.

Suppose that $y_i(x)$, $u_s(x)$ is an admissible arc for $\mathfrak{B}$ which is without corners, and which satisfies with multipliers $\lambda_0 = 1$, $\lambda_\kappa(x)$ $(\kappa = 1, \cdots, m + q)$ and constants e_μ, e_s' the multiplier rule, the Weierstrass condition II_N, the non-singularity condition, and IV_*'. Then $\mathfrak{E}$: $y_i(x)$, $u_s(x)$, $\lambda_0 = 1$, $\lambda_\kappa(x)$ is an extremal for $\mathfrak{B}$. In particular, the multiplier rule implies that $\lambda_{m+s}(x)$ $(s = 1, \cdots, q)$ are constant on x_1x_2 and that $e_s' = \lambda_{m+s}$. Corresponding to $\mathfrak{E}$ we denote by E the set $y_i(x)$, $\lambda_0 = 1$, $\lambda_\kappa(x)$ $(x_1 \leqq x \leqq x_2)$; such a set will be termed an extremal for B. If we set

$$F[x, y, r, \lambda] = \lambda_0 f[x, y, r] + \lambda_\alpha \phi_\alpha[x, y, r] + \lambda_{m+s} h_s[x, y, r],$$

the multiplier rule implies that along E

$$\frac{dF_{r_i}}{dx} - F_{y_i} = 0, \qquad \phi_\alpha = 0,$$

and the relation

$$[(F - y_i' F_{r_i})\, dx + F_{r_i}\, dy_i]_1^2 + dg + e_\mu\, d\psi_\mu + \lambda_{m+s}\, d\chi_s = 0$$

holds for every choice of the differential dx_1, dy_{i1}, dx_2, dy_{i2}. Since problem $\mathfrak{B}$ does not involve the variables u_s explicitly and the differential equations (2.2′) contain the u_s' only linearly, it readily follows that II_N is equivalent to the condition that there exist a neighborhood N of the elements of E in $[x, y, r, \lambda]$-space such that if $[x, y, r, \lambda]$ is in N and $\phi_\alpha[x, y, r] = 0$, $\phi_\alpha[x, y, \bar{r}] = 0$, then

$$\mathcal{E}[x, y, r, \lambda; \bar{r}] \equiv F[x, y, \bar{r}, \lambda] - F[x, y, r, \lambda] - (\bar{r}_i - r_i) F_{r_i}[x, y, r, \lambda] \tag{2.7}$$

is non-negative. The non-singularity condition for $\mathfrak{B}$ is reducible to the non-singularity of the matrix

$$\begin{Vmatrix} F_{r_i r_j} & \phi_{\beta r_i} \\ \phi_{\alpha r_j} & 0_{\alpha\beta} \end{Vmatrix}$$

along the elements of E. Finally, IV_*' may be written as the condition that along E the second variation

$$J_2 = 2\gamma[\xi_1, \eta(x_1), \xi_2, \eta(x_2)] + \int_{x_1}^{x_2} 2\omega[x, \eta, \eta']\, dx \tag{2.8}$$

be positive for all non-identically vanishing sets $[\xi_1, \xi_2, \eta_i(x)]$ satisfying along E the equations of variation

$$\Phi_\alpha[x, \eta, \eta'] = 0, \qquad \Psi_\mu[\xi_1, \eta(x_1), \xi_2, \eta(x_2)] = 0,$$

$$X_s[\xi_1, \eta(x_1), \xi_2, \eta(x_2)] + \int_{x_1}^{x_2} (h_{sr_i}\eta_i' + h_{sy_i}\eta_i)\, dx = 0.$$

In (2.8)

$$2\omega \equiv F_{r_ir_j}\eta_i'\eta_j' + 2F_{r_iy_j}\eta_i'\eta_j + F_{y_iy_j}\eta_i\eta_j ,$$

and 2γ is a quadratic form in its arguments whose explicit form will not be given.[3]

It is well known that conditions II_N and non-singularity imply the strengthened Clebsch condition III'; that is, at the elements $[x, y(x), y'(x), \lambda(x)]$ $(x_1 \leqq x \leqq x_2)$ of E the quadratic form $F_{r_ir_j}\pi_i\pi_j$ is positive for all sets $(\pi_i) \neq (0_i)$ satisfying $\phi_{\alpha r_j}\pi_j = 0$ $(\alpha = 1, \cdots, m)$. It has also been pointed out recently by Hestenes and the author (see [6]) that conditions II_N and non-singularity imply the existence of a neighborhood of the elements of E in $[x, y, r, \lambda]$-space such that for $[x, y, r, \lambda]$ in this neighborhood, $\phi_\alpha[x, y, r] = 0$, $\phi_\alpha[x, y, \tilde{r}] = 0$, and $(\tilde{r}_i) \neq (r_i)$, the $\mathfrak{E}$-function of (2.7) is positive. This neighborhood in general may be smaller than N, but for simplicity *we shall suppose that N refers to such a restricted neighborhood*; this strengthened condition will be denoted, as usual, by II_N'.

If $C: Y_i(x)$ $(X_1 \leqq x \leqq X_2)$ is admissible for the problem B, we shall denote by $\mathfrak{C}$ the set

$$\mathfrak{C}: Y_i(x),\ U_s(x) = \int_x^{X_2} h_s[t, Y(t), Y'(t)]\, dt \qquad (X_1 \leqq x \leqq X_2),$$

which defines an admissible arc for the problem $\mathfrak{B}$. The usual sufficiency theorem for the problem of Bolza gives the following result:

THEOREM 2.1. *Suppose that an admissible arc* (2.1) *for* B *is without corners, and $E: y_i(x)$, $\lambda_0 = 1$, $\lambda_\kappa(x)$ $(x_1 \leqq x \leqq x_2)$ satisfies with constants e_μ the multiplier rule, II_N, non-singularity, and IV_*'. If $\mathfrak{E}$ denotes the corresponding set $\mathfrak{E}: y_i(x)$, $u_s(x)$, $\lambda_0 = 1$, $\lambda_\kappa(x)$, then there exists a neighborhood $\mathfrak{F}$ of $\mathfrak{E}$ in xyu-space, and a neighborhood $\mathfrak{M}$ of the ends of $\mathfrak{E}$ in $[x_1, y_{i1}, u_{s1}, x_2, y_{i2}, u_{s2}]$-space such that $J[\mathfrak{C}] > J[\mathfrak{E}]$ for every admissible arc $\mathfrak{C}$ in $\mathfrak{F}$ with ends in $\mathfrak{M}$ and not identical with $\mathfrak{E}$.*

The central purpose of this paper is to prove the following stronger theorem.

THEOREM 2.2. *Under the hypotheses of Theorem* 2.1 *there exist a neighborhood $\mathcal{F}$ of E in xy-space and a neighborhood $\mathcal{M}$ of the ends of E in $[x_1, y_{i1}, x_2, y_{i2}]$-space such that $J[C] > J[E]$ for every admissible arc C in $\mathcal{F}$ with ends in $\mathcal{M}$ and not identical with E.*

The passage from the result of Theorem 2.1 to that of Theorem 2.2 will be effected by the use of the Lindeberg theorem of §4. Preliminary to the derivation of this latter result we shall, however, discuss in the following section a problem B* which is entirely equivalent to B, and for which stronger forms of the Clebsch and Weierstrass condition are satisfied.

[3] For a discussion of the second variation, as well as its explicit form, the reader is referred to Bliss [1], pp. 68–71; see also Hestenes [3], p. 797, or Reid [10], p. 665.

3. An associated minimum problem. Let

$$f^*[x, y, r] = f[x, y, r] + \tfrac{1}{2}l(x)\phi_\alpha[x, y, r]\phi_\alpha[x, y, r],$$

where $l(x)$ is a given function of class C^2 on x_1x_2, and denote by B^* the problem in which the expressions ϕ_α, ψ_μ, χ_s, g, h_s are as in B, but f is replaced by f^*. Clearly an arc (2.1) affords a minimum for B if and only if it affords a minimum for B^*. If

$$F^*[x, y, r, \lambda] = \lambda_0 f^*[x, y, r] + \lambda_\alpha\phi_\alpha[x, y, r] + \lambda_{m+s}h_s[x, y, r],$$

the functions F and F^*, together with their first order partial derivatives, are identical at a set for which $\phi_\alpha[x, y, r] = 0$. Hence $E: y_i(x)$, $\lambda_0 = 1$, $\lambda_s(x)$ satisfies with constants e_μ the multiplier rule for B if and only if E satisfies with the same constants e_μ the multiplier rule for B^*. Along E, $F^*_{r_ir_j} = F_{r_ir_j} + l(x)\phi_{\alpha r_i}\phi_{\alpha r_j}$; moreover, if J_2^* denotes the second variation along E for B, then

$$J_2^* = J_2 + \int_{x_1}^{x_2} l(x)\Phi_\alpha[x, \eta, \eta']\Phi_\alpha[x, \eta, \eta']\, dx.$$

Consequently, each of the following conditions is satisfied for both the problems B and B^* whenever it is satisfied for one of them: II, II_N, II'_N, non-singularity, III, III', IV_*, IV'_*.[4]

It will now be proved that whenever E is an extremal for B satisfying III' and II'_N, the function $l(x)$ may be so chosen that the corresponding problem B^* satisfies certain stronger forms of these conditions.

THEOREM 3.1. *If E is an extremal for* B *satisfying* III', *there exists a constant l such that along E the corresponding quadratic form*

$$F^*_{r_ir_j}\pi_i\pi_j \equiv (F_{r_ir_j} + l\,\phi_{\alpha r_i}\phi_{\alpha r_j})\pi_i\pi_j \tag{3.1}$$

for B^* *is positive definite.*

Condition III' is equivalent to the assumption that the quadratic form $F_{r_ir_j}\pi_i\pi_j$ is positive for all sets $(\pi_i) \neq (0_i)$ which give the quadratic form $(\phi_{\alpha r_i}\phi_{\alpha r_j})\pi_i\pi_j$ the value zero. Since the last quadratic form is positive semi-definite, one may prove the above theorem in a direct and quite elementary fashion. This theorem is, however, a ready consequence of a theorem on quadratic forms proved by the author.[5] For, as there proved, the determinant $P(\sigma) = |F_{r_ir_j} - \sigma\phi_{\alpha r_i}\phi_{\alpha r_j}|$ is a polynomial of degree m in σ and has only real roots. Moreover, since the form $(\phi_{\alpha r_i}\phi_{\alpha r_j})\pi_i\pi_j$ is positive semi-definite, it follows from the proof of Reid [12] that if $-l(x)$ is less than the smallest zero of $P(\sigma)$ for the arguments $[x, y(x), y'(x), \lambda(x)]$ of E, then $(F_{r_ir_j} + l(x)\phi_{\alpha r_i}\phi_{\alpha r_j})\pi_i\pi_j$

[4] For the explicit definition and discussion of the conditions in this list that have not been defined above, the reader is referred to Bliss [1], Hestenes [3], or Reid [10].

[5] See Reid [12]; also A. A. Albert, Bulletin of the American Mathematical Society, vol. 44(1938), pp. 250-253.

is a positive definite quadratic form. In view of the continuity of the coefficients of $P(\sigma)$ as functions of x, there exists a constant l such that for $x_1 \leqq x \leqq x_2$ the form (3.1) is positive definite along E. In other words, if E satisfies III' for B there is a suitably chosen problem B* for which the strengthened Clebsch condition holds without the restriction $\phi_{\alpha r_j}\pi_j = 0$. This new condition will be denoted by III*.

The Weierstrass $\mathcal{E}$-function for the problem B* will be denoted by $\mathcal{E}^*[x, y, r, \lambda; \bar{r}]$. In the future we shall be concerned only with a problem B* for which the function $l(x)$ is a constant value that satisfies the condition of Theorem 3.1. The following corollary is an immediate consequence of a simple continuity argument and an application of Taylor's formula.

COROLLARY. *If E is an extremal for* B *satisfying condition* III*, *there exists a positive constant τ_0 and a neighborhood $\mathfrak{N}_0$ of the elements of E in $[x, y, r, \lambda]$-space such that if $[x, y, r, \lambda]$ and $[x, y, \bar{r}, \lambda]$ are in $\mathfrak{N}_0$, then*[6]

$$\mathcal{E}^*[x, y, r, \lambda; \bar{r}] \geqq \tau_0 \| \bar{r} - r \|^2. \tag{3.2}$$

THEOREM 3.2. *If E: $y_i(x)$, $\lambda_0 = 1$, $\lambda_\kappa(x)$ $(x_1 \leqq x \leqq x_2)$ is an extremal for* B *satisfying conditions* II'_N, III' *and l is chosen as in Theorem* 3.1, *there exists a neighborhood $\mathfrak{N}$ of the elements of E in $[x, y, r, \lambda]$-space such that for $[x, y, r, \lambda]$ in $\mathfrak{N}$, $[x, y, r] \neq [x, y, \bar{r}]$, and $\phi_\alpha[x, y, \bar{r}] = 0$, we have $\mathcal{E}^*[x, y, r, \lambda; \bar{r}] > 0$.*

The conclusion of this theorem is a stronger condition than II'_N since it does not require that $\phi_\alpha[x, y, r] = 0$. This latter condition will be denoted by $\mathrm{II}^*_{\mathfrak{N}}$.

The proof of this theorem is essentially the same as the proof of Lemma 4.3 of Reid [10]. In view of the non-singularity of B, and hence of B*, there exists a neighborhood $\mathfrak{N}$ of the elements of E in $[x, y, r, \lambda]$-space and a constant δ such that for $[x, y, r, \lambda]$ in $\mathfrak{N}$ the system

$$\begin{aligned} &F^*_{r_i}[x, y, r + z, \lambda + \nu] - F^*_{r_i}[x, y, r, \lambda] = 0 \quad (i = 1, \cdots, n), \\ &\phi_\alpha[x, y, r + z] = 0, \qquad \nu_{m+s} = 0 \qquad (\alpha = 1, \cdots, m;\ s = 1, \cdots, q) \end{aligned} \tag{3.3}$$

has a unique solution $z_i = z_i[x, y, r, \lambda]$, $\nu_\alpha = \nu_\alpha[x, y, r, \lambda]$ satisfying $z_i z_i + \nu_\alpha\nu_\alpha < \delta$. We may obviously restrict $\mathfrak{N}$ so that for $[x, y, r, \lambda]$ in $\mathfrak{N}$ the corresponding set $[x, y, r + z, \lambda + \nu]$ is in the neighborhood N and also in the neighborhood $\mathfrak{N}_0$ of the above corollary. Then for $[x, y, r, \lambda]$ in $\mathfrak{N}$ and $\bar{r}$ such that $\phi_\alpha[x, y, \bar{r}] = 0$, it follows from (3.3) that

$$\mathcal{E}^*[x, y, r, \lambda; \bar{r}] = \mathcal{E}^*[x, y, r + z, \lambda + \nu; \bar{r}] + \mathcal{E}^*[x, y, r, \lambda; r + z].$$

Since $[x, y, r + z, \lambda + \nu]$ is in N and $\phi_\alpha[x, y, r + z] = 0$, condition II'_N implies $\mathcal{E}^*[x, y, r + z, \lambda + \nu; \bar{r}] \geqq 0$, the equality holding only if $\bar{r}_i = r_i + z_i$. Moreover, as $[x, y, r, \lambda]$ and $[x, y, r + z, \lambda]$ are in $\mathfrak{N}_0$, we have $\mathcal{E}^*[x, y, r, \lambda; r + z] \geqq 0$, the equality holding only if $z_i = 0_i$. The conclusion of Theorem 3.2 is an immediate consequence of these results.

[6] If $z = (z_1, \cdots, z_k)$, $\| z \|$ is used to denote the positive square root of $z_1^2 + \cdots + z_k^2$.

We shall denote by $\mathcal{E}_{h_s}$ the Weierstrass $\mathcal{E}$-function for h_s, that is,

$$\mathcal{E}_{h_s}[x, y, r; \bar{r}] = h_s[x, y, \bar{r}] - h_s[x, y, r] - (\bar{r}_i - r_i)h_{sr_i}[x, y, r],$$

and represent by $\mathcal{E}_h$ the set $(\mathcal{E}_{h_s})$ $(s = 1, \cdots, q)$. The following corollary to the above theorem will be used in the proof of the Lindeberg theorem of the next section.

Corollary. *If E satisfies the hypotheses of Theorem 3.2, there exists a positive constant k and a neighborhood $\mathfrak{N}_1$ of the elements of E in $[x, y, r, \lambda]$-space such that if $[x, y, r, \lambda]$ is in $\mathfrak{N}_1$ and $\phi_\alpha[x, y, \bar{r}] = 0$, then*

$$\| \mathcal{E}_h[x, y, r; \bar{r}] \| \leqq k\mathcal{E}^*[x, y, r, \lambda; \bar{r}]. \tag{3.4}$$

For let $\mathfrak{N}_1$ be a neighborhood of the elements of E in $[x, y, r, \lambda]$-space such that there is an associated constant $d > 0$ satisfying the condition that if $[x, y, r, \lambda]$ is in $\mathfrak{N}_1$ and $\nu_\kappa\nu_\kappa \leqq d$, $\nu_\alpha = 0$ $(\alpha = 1, \cdots, m)$, then $[x, y, r, \lambda \pm \nu]$ is in the neighborhood $\mathfrak{N}$ of Theorem 3.2. For $[x, y, r, \lambda]$ in $\mathfrak{N}_1$ and $\phi_\alpha[x, y, \bar{r}] = 0$ we then have

$$0 \leqq \mathcal{E}^*[x, y, r, \lambda \pm \nu; \bar{r}] = \mathcal{E}^*[x, y, r, \lambda; \bar{r}] \pm \nu_{m+s}\mathcal{E}_{h_s}[x, y, r; \bar{r}],$$

that is,

$$| \nu_{m+s}\mathcal{E}_{h_s}[x, y, r; \bar{r}] | \leqq \mathcal{E}^*[x, y, r, \lambda; \bar{r}]$$

for all sets ν_{m+s} satisfying $\nu_{m+s}\nu_{m+s} \leqq d$. It is then a consequence of Cauchy's inequality that (3.4) holds for $k = d^{-\frac{1}{2}}$.

Using the above results, one may readily prove for $\mathcal{E}^*$ an inequality corresponding to that established for $\mathcal{E}$ by Reid ([10], Theorems 4.1 and 4.2). Instead of the function $R[t]$ there employed, we shall use here the function $\mathfrak{R}[t] = t^2/(t + 1)$ for $t \geqq 0$. In some respects this function is simpler than the $R[t]$ previously used. In particular, it is easily seen that for $t \geqq 0$

$$\mathfrak{R}[t] \leqq \min (t, t^2) \leqq 2\mathfrak{R}[t]. \tag{3.5}$$

Theorem 3.3. *If E satisfies the hypotheses of Theorem 3.2, there exists a positive constant τ and a neighborhood $\mathfrak{N}_2$ of the elements of E in $[x, y, r, \lambda]$-space such that for $[x, y, r, \lambda]$ in $\mathfrak{N}_2$ and $\phi_\alpha[x, y, \bar{r}] = 0$ we have*

$$\mathcal{E}^*[x, y, r, \lambda; \bar{r}] \geqq \tau\mathfrak{R}[\| \bar{r} - r \|]. \tag{3.6}$$

Let $\mathfrak{N}_2$ be a bounded neighborhood of the elements of E in $[x, y, r, \lambda]$-space with the property that there exists a constant $d > 0$ such that if $[x, y, r, \lambda]$ is in $\mathfrak{N}_2$ and $\| v \| \leqq 2d$, then $[x, y, r + v, \lambda]$ is in the neighborhood $\mathfrak{N}_0$ of the corollary to Theorem 3.1 and also in $\mathfrak{N}$ determined by Theorem 3.2. If $[x, y, r, \lambda]$ is in $\mathfrak{N}_2$ and $\| \bar{r} - r \| \leqq 2d$, it is then a consequence of (3.5) and the corollary to Theorem 3.1 that (3.6) holds for $\tau = \tau_0$.

Now suppose that $[x, y, r, \lambda]$ is in $\mathfrak{N}_2$, the set $[x, y, \bar{r}]$ satisfies $\phi_\alpha[x, y, \bar{r}] = 0$,

and $\| \bar{r} - r \| \geqq 2d$. Let $r^* = r + \theta(\bar{r} - r)$, where θ is so chosen that $\theta \| \bar{r} - r \| = d$; clearly $0 < \theta \leqq \frac{1}{2}$. If[7]

$$T[x, y, r, \lambda; \rho] = F^*[x, y, r, \lambda] + (\rho_i - r_i)F^*_{r_i}[x, y, r, \lambda],$$

we have

$$T[x, y, r^*, \lambda; r^*] - T[x, y, r, \lambda; r^*] = \mathcal{E}^*[x, y, r, \lambda; r^*] > 0,$$

$$T[x, y, r^*, \lambda; r] - T[x, y, r, \lambda; r] = -\mathcal{E}^*[x, y, r^*, \lambda; r] < 0.$$

Consequently, there exists a value $\tilde{r} = r + \sigma(\bar{r} - r)$, $0 < \sigma < \theta$, such that $T[x, y, r^*, \lambda; \tilde{r}] = T[x, y, r, \lambda; \tilde{r}]$, and this relation implies

$$\mathcal{E}^*[x, y, r, \lambda; \bar{r}] = \mathcal{E}^*[x, y, r^*, \lambda; \bar{r}] + \frac{1 - \sigma}{\sigma} \mathcal{E}^*[x, y, r^*, \lambda; r]. \tag{3.7}$$

Since $[x, y, r^*, \lambda]$ is in $\mathfrak{N}$, $\mathcal{E}^*[x, y, r^*, \lambda; \bar{r}] \geqq 0$; as $0 < \sigma < \theta \leqq \frac{1}{2}$, $(1 - \sigma)/\sigma > (1 - \theta)/\theta \geqq 1/(2\theta) = (1/(2d)) \| \bar{r} - r \|$. Hence if τ_1 is such that $\mathcal{E}^*[x, y, r^*, \lambda; r] \geqq \tau_1$ for all $[x, y, r^*, \lambda]$ in $\mathfrak{N}$, $\| r^* - r \| = d$, it follows from (3.7) and property (3.5) that

$$\mathcal{E}^*[x, y, r, \lambda; \bar{r}] \geqq (\tau_1/(2d))\mathfrak{R}[\| \bar{r} - r \|].$$

It has been proved, therefore, that with $\mathfrak{N}_2$ chosen as indicated above the inequality (3.6) holds for $\tau = \min(\tau_0, \tau_1/(2d))$. The simplicity of this argument in comparison to the proof of Theorem 4.2 of Reid [10] results from the above Theorem 3.2 for the suitably chosen problem B*. Since in case $\phi_\alpha[x, y, r] = 0$, $\phi_\alpha[x, y, \bar{r}] = 0$ we have $\mathcal{E}^*[x, y, r, \lambda; \bar{r}] = \mathcal{E}[x, y, r, \lambda; \bar{r}]$, the results of Theorems 4.1 and 4.2 of Reid [10] are included as a very special case of the above theorem. If B is an ordinary problem of Bolza not involving the isoperimetric conditions (2.4), and one applies the expansion method of proof to a suitably chosen problem B* instead of to the original problem B, the proof of Theorem 6.1 of Reid [10] is materially simplified. In particular, Theorem 4.3 of that paper is no longer needed, since an even stronger form of its conclusion is an immediate consequence of the above Theorem 3.3.

4. A Lindeberg theorem. In the proof of the results of this section use is made of the following lemma.

LEMMA 4.1. *If $z_t(x)$ $(t = 1, \cdots, n)$ are absolutely continuous functions on $X_1 \leqq x \leqq X_2$ and $\| z \| \leqq \delta$ on this interval, then*

$$\int_{X_1}^{X_2} \| z \| \cdot \| z' \| \, dx \leqq d_1 \int_{X_1}^{X_2} \mathfrak{R}[\| z' \|] \, dx + \| z(X_1) \|^2, \tag{4.1}$$

[7] For fixed values $[x, y, \lambda]$, $\zeta = T[x, y, r, \lambda; \rho]$ is the equation in (ρ, ζ)-space of the tangent plane to the surface $\zeta = F^*[x, y, \lambda, \rho]$ at $\rho = r$. The following proof is essentially the same as the geometric proof given by Tonelli ([14], vol. 1, pp. 351–353) of a corresponding inequality for the plane problem of the calculus of variations.

$$\int_{X_1}^{X_2} \| z \|^2 \, dx \leqq d_2 \int_{X_1}^{X_2} \mathfrak{R}[\| z' \|] \, dx + \| z(X_1) \|^2, \tag{4.2}$$

where $d_1 = 2 \max(1, 2\delta) \max(1, X_2 - X_1)$ *and* $d_2 = (X_2 - X_1) d_1$.

This lemma is established in a manner analogous to the proof of Theorem 5.1 of Reid [10]; use is made of relation (3.5) and the following properties of the positive convex function $\mathfrak{R}[t]$ $(t \geqq 0)$ here employed:

P_1 : If $a \geqq 0$, $a \min(1, a)\mathfrak{R}[t] \leqq \mathfrak{R}[at] \leqq a \max(1, a)\mathfrak{R}[t]$;

P_2 : If $t_1 \geqq 0$, $t_2 \geqq 0$, $\mathfrak{R}[t_1 + t_2] \leqq 2(\mathfrak{R}[t_1] + \mathfrak{R}[t_2])$.

Property P_1 is an immediate consequence of the relation $\mathfrak{R}[at]/\mathfrak{R}[t] = a^2[(t + 1)/(at + 1)]$; P_2 follows by the use of an elementary inequality which gives $\mathfrak{R}[t_1 + t_2] \leqq 2(t_1^2 + t_2^2)/(t_1 + t_2 + 1) \leqq 2[t_1^2/(t_1 + 1) + t_2^2/(t_2 + 1)]$.

Suppose $E: y_i(x)$, $\lambda_0 = 1$, $\lambda_\epsilon(x)$ $(x_1 \leqq x \leqq x_2)$ is an extremal for B which satisfies II_N and the non-singularity condition. As pointed out in §2, these conditions imply III' and II'_N. In view of the non-singularity condition there exists an $\epsilon_0 > 0$ such that for $| X_\nu - x_\nu | \leqq \epsilon_0$ $(\nu = 1, 2)$ the elements of E are defined and of class C^2 on X_1X_2. The corresponding extremal $\mathfrak{E}$: $y_i(x)$, $u_s(x)$, $\lambda_0 = 1$, $\lambda_\epsilon(x)$ for $\mathfrak{B}$ clearly has its elements also defined and of class C^2 on X_1X_2. In the proof of an effective Lindeberg theorem for the problem B we shall use the following notation. An arc C: $Y_i(x)$ $(X_1 \leqq x \leqq X_2)$ will be said to lie in $(E)_\epsilon$ if $\| Y(x) - y(x) \| \leqq \epsilon$ $(X_1 \leqq x \leqq X_2)$ and $| X_\nu - x_\nu | \leqq \epsilon$ $(\nu = 1, 2)$; similarly, an arc $\mathfrak{C}$: $Y_i(x)$, $U_s(x)$ $(X_1 \leqq x \leqq X_2)$ will be said to lie in $(\mathfrak{E})_\epsilon$ if $\| Y(x) - y(x) \| \leqq \epsilon$, $\| U(x) - u(x) \| \leqq \epsilon$ $(X_1 \leqq x \leqq X_2)$ and $| X_\nu - x_\nu | \leqq \epsilon$. In the definitions of $(E)_\epsilon$ and $(\mathfrak{E})_\epsilon$ nothing is prescribed concerning the continuity and differentiability of the functions $Y_i(x)$, $U_s(x)$. It will be understood in the following that the values of ϵ used do not exceed ϵ_0.

THEOREM 4.1. *If* $E: y_i(x)$, $\lambda_0 = 1$, $\lambda_\epsilon(x)$ $(x_1 \leqq x \leqq x_2)$ *is an extremal for* B *which satisfies conditions* II_N *and non-singularity, then corresponding to a given* $\eta > 0$ *there exist constants* $\rho > 0$, $\epsilon > 0$ *such that if* C: $Y_i(x)$ $(X_1 \leqq x \leqq X_2)$ *is an admissible arc for* B *which is in* $(E)_\epsilon$, *while*

$$\int_{X_1}^{X_2} \mathcal{E}^*[x, Y(x), y'(x), \lambda(x); Y'(x)] \, dx \leqq \rho, \tag{4.3}$$

then the corresponding arc $\mathfrak{C}$: $Y_i(x)$, $U_s(x)$ $(X_1 \leqq x \leqq X_2)$ *is in* $(\mathfrak{E})_\eta$.

Let $\epsilon_1 > 0$ be such that if C is in $(E)_{\epsilon_1}$, then the set $[x, Y(x), y'(x), \lambda(x)]$ $(X_1 \leqq x \leqq X_2)$ is in $\mathfrak{N}_1$ defined by the corollary to Theorem 3.2, and also in the neighborhood $\mathfrak{N}_2$ of Theorem 3.3. It is supposed, moreover, that ϵ_1 is such that if C is in $(E)_{\epsilon_1}$ then $[x, Y(x), y'(x)]$ $(X_1 \leqq x \leqq X_2)$ is interior to the region R; in particular, on the set of values $[x, Y(x), y'(x)]$ the functions h_s, together with their partial derivatives of the first three orders, are uniformly bounded with respect to all C: $Y_i(x)$ $(X_1 \leqq x \leqq X_2)$ in $(E)_{\epsilon_1}$.

Now

$$
\begin{aligned}
U_s(x) - u_s(x) &= \int_x^{X_2} h_s[x, Y, Y']\,dx - \int_x^{x_2} h_s[x, y, y']\,dx \\
&= \int_x^{X_2} \mathcal{E}_{h_s}[x, Y, y'; Y']\,dx + \int_{x_2}^{X_2} h_s[x, y, y']\,dx \\
&\quad + \int_x^{X_2} (h_s[x, Y, y'] - h_s[x, y, y'] + (Y_i' - y_i')h_{sr_i}[x, Y, y'])\,dx \\
&= \int_x^{X_2} \mathcal{E}_{h_s}[x, Y, y'; Y']\,dx + I_{1s} + I_{2s},
\end{aligned}
\tag{4.4}
$$

where

$$
\begin{aligned}
I_{1s} &= \int_{x_2}^{X_2} h_s[x, y, y']\,dx + (Y_i - y_i)h_{sr_i}[x, y, y']\,\Big|_{x=x}^{x=X_2} \\
&\quad + \int_x^{X_2} (h_s[x, Y, y'] - h_s[x, y, y'] - (Y_i - y_i)\,dh_{sr_i}[x,y,y']/dx)\,dx, \\
I_{2s} &= \int_x^{X_2} (Y_i' - y_i')(h_{sr_i}[x, Y, y'] - h_{sr_i}[x, y, y'])\,dx.
\end{aligned}
$$

Clearly the norm of the set $(U_s(x) - u_s(x))$ $(s = 1, \cdots, q)$ satisfies the inequality

$$
\| U(x) - u(x) \| \leqq \left\| \int_x^{X_2} \mathcal{E}_h[x, Y, y'; Y']\,dx \right\| + \| I_1 \| + \| I_2 \|. \tag{4.5}
$$

If C is such that $[x, Y(x)]$ is in $(E)_\epsilon$ and $0 < \epsilon < \epsilon_1$, we have

$$
\begin{aligned}
\left\| \int_x^{X_2} \mathcal{E}_h[x, Y, y'; Y']\,dx \right\| &\leqq \int_{X_1}^{X_2} \| \mathcal{E}_h[x, Y, y'; Y'] \|\,dx \\
&\leqq k \int_{X_1}^{X_2} \mathcal{E}^*[x, Y, y', \lambda; Y']\,dx,
\end{aligned}
\tag{4.6a}
$$

the last inequality being a consequence of the corollary to Theorem 3.2. *If we use c as a generic constant,* it is a consequence of the elements of C being in $(E)_\epsilon$ that

$$
\| I_1 \| \leqq c\epsilon. \tag{4.6b}
$$

Moreover, since for $0 < \epsilon < \epsilon_1$ the neighborhood $(E)_\epsilon$ is interior to $\mathfrak{N}_2$ determined by Theorem 3.3, one may prove by the use of the inequality (4.1) that

$$
\begin{aligned}
\| I_2 \| &\leqq c \int_{X_1}^{X_2} \| Y - y \| \cdot \| Y' - y' \|\,dx \\
&\leqq cd_1 \left[\int_{X_1}^{X_2} \Re[\| Y' - y' \|]\,dx + \| Y(X_1) - y(X_1) \|^2 \right] \\
&\leqq cd_1 \left[\frac{1}{\tau} \int_{X_1}^{X_2} \mathcal{E}^*[x, Y, y', \lambda; Y']\,dx + \epsilon^2 \right].
\end{aligned}
\tag{4.6c}
$$

The constant d_1 appearing in (4.6c) depends upon ϵ and $X_2 - X_1$, but is obviously bounded since $| X_\nu - x_\nu | \leqq \epsilon$ and $0 < \epsilon < \epsilon_1$. In the above inequalities explicit values of the constants c could be given in terms of the bounds of the functions h_s and certain of their partial derivatives.

It follows readily from the inequalities (4.6) that positive constants ρ, ϵ may be so chosen that for an admissible arc C in $(E)_\epsilon$ and satisfying (4.3) we have $\| U(x) - u(x) \| < \eta$ on X_1X_2, where η is any preassigned positive value. In particular, if such an ϵ is chosen not to exceed the given η, the arc $\mathfrak{C}$: $Y_i(x)$, $U_s(x)$, $(X_1 \leqq x \leqq X_2)$ is in $(\mathfrak{E})_\eta$.

THEOREM 4.2. *If E: $y_i(x)$, $\lambda_0 = 1$, $\lambda_\kappa(x)$ $(x_1 \leqq x \leqq x_2)$ is an extremal for* B *which satisfies conditions* II_N *and non-singularity, and ρ is a given positive constant, then there exists an $\epsilon' > 0$ such that $J[C] - J[E] \geqq \frac{1}{2}\rho$ for every admissible arc C: $Y_i(x)$ $(X_1 \leqq x \leqq X_2)$ whose elements are in $(E)_{\epsilon'}$ and for which*

$$\int_{X_1}^{X_2} \mathcal{E}^*[x, Y(x), y'(x), \lambda(x); Y'(x)]\,dx > \rho. \tag{4.7}$$

Let $\epsilon_2 > 0$ be such that if C: $Y_i(x)$ $(X_1 \leqq x \leqq X_2)$ is an arc with elements in $(E)_{\epsilon_2}$, then the set $[x, Y(x), y'(x), \lambda(x)]$ is in the neighborhood $\mathfrak{N}_2$ of Theorem 3.3. It is supposed, moreover, that ϵ_2 is such that if C is in $(E)_{\epsilon_2}$, then $[x, Y(x), y'(x)]$ $(X_1 \leqq x \leqq X_2)$ is in the region R, and the set $[X_1, Y_i(X_1), X_2, Y_i(X_2)]$ is in R_1. In particular, the function F^* and its partial derivatives of the first three orders are bounded on the set of values $[x, Y(x), y'(x), \lambda(x)]$ uniformly with respect to all arcs C: $Y_i(x)$ $(X_1 \leqq x \leqq X_2)$ in $(E)_{\epsilon_2}$; similarly, the function g and its partial derivatives of the first two orders are bounded on the set of values $[X_1, Y_i(X_1), X_2, Y_i(X_2)]$ uniformly with respect to such arcs C.

For an admissible arc C in $(E)_{\epsilon'}$ $(0 < \epsilon' \leqq \epsilon_2)$ we then have

$$\begin{aligned} J[C] - J[E] &= \Delta g + \int_{X_1}^{X_2} f^*[x, Y, Y']\,dx - \int_{x_1}^{x_2} f^*[x, y, y']\,dx \\ &= \Delta g + \int_{X_1}^{X_2} F^*[x, Y, Y', \lambda]\,dx - \int_{x_1}^{x_2} F^*[x, y, y', \lambda]\,dx \\ &= \Delta g + \int_{X_1}^{X_2} \mathcal{E}^*[x, Y, y', \lambda; Y']\,dx + J^1 + J^2, \end{aligned} \tag{4.8}$$

where

$$\begin{aligned} J^1 = \int_{X_1}^{x_1} F^*[x, y, y', \lambda]\,dx &+ \int_{x_2}^{X_2} F^*[x, y, y', \lambda]\,dx + (Y_i - y_i)F^*_{r_i}[x, y, y', \lambda]\,\Big|_{x=X_1}^{x=X_2} \\ &+ \int_{X_1}^{X_2} \{F^*[x, Y, y', \lambda] - F^*[x, y, y', \lambda] - (Y_i - y_i)F^*_{y_i}[x, y, y', \lambda]\}\,dx, \end{aligned}$$

$$J^2 = \int_{X_1}^{X_2} (Y_i' - y_i')\{F^*_{r_i}[x, Y, y', \lambda] - F^*_{r_i}[x, y, y', \lambda]\}\,dx.$$

The last term in J^1 is written involving $F^*_{y_i}$ since along E, $F^*_{y_i} = dF^*_{r_i}/dx$. In the expression (4.8), $\Delta g = g[X_1, Y(X_1), X_2, Y(X_2)] - g[x_1, y(x_1), x_2, y(x_2)]$. *Again using c as a generic constant*, we have

$$
\begin{aligned}
&(4.9\text{a}) \qquad | \Delta g | \leqq c\epsilon', \qquad | J^1 | \leqq c\epsilon', \\
&\qquad | J^2 | \leqq c \int_{X_1}^{X_2} \| Y - y \| \cdot \| Y' - y' \| \, dx \\
&\qquad\qquad \leqq c\epsilon' \int_{X_1}^{X_2} \| Y' - y' \| \, dx \\
&\qquad\qquad \leqq c\epsilon' \int_{X_1}^{X_2} (1 + \mathfrak{R}[\| Y' - y' \|]) \, dx \\
&(4.9\text{b}) \qquad\qquad \leqq c\epsilon' \left[(X_2 - X_1) + \frac{1}{\tau} \int_{X_1}^{X_2} \mathcal{E}^*[x, Y, y', \lambda; Y'] \, dx \right].
\end{aligned}
$$

In the above inequalities explicit values of the constants c could be given in terms of the bounds of the functions F^*, g, and certain of their partial derivatives. The third of the above inequalities for $| J^2 |$ follows from the preceding by the easily established relation $t \leqq 1 + \mathfrak{R}[t]$.

By the use of inequalities (4.9) it follows readily that for ρ a preassigned constant there exists an $\epsilon' > 0$ such that if C is admissible for B and has its elements in $(E)_{\epsilon'}$, and if for C inequality (4.7) is satisfied, then $J[C] - J[E] > \frac{1}{2}\rho$. Thus the theorem is proved.

As stated in the introduction, for non-parametric problems neither the result obtained by Lindeberg, nor any one of the extended forms of the Lindeberg theorem established by Levi and Tonelli, is effective in the proof of Theorem 2.2. Each of the referred to results gives an analogue of Theorem 4.2; it is the analogue of Theorem 4.1 that is incapable of proof by these forms of the Lindeberg theorem. For brevity, we shall not state explicitly any one of these results. For one familiar with these conditions, however, the following simple example illustrates why they are not applicable to the proof of a result analogous to that of Theorem 4.1.

Consider the problem of minimizing the integral

$$(4.10) \qquad J = \int_0^2 y'^2 \, dx$$

in the class of arcs $y = y(x)$ satisfying the end-conditions

$$(4.11) \qquad y(0) = 0 = y(2)$$

and the isoperimetric condition

$$(4.12) \qquad \int_0^2 (1 - x) y'^2 \, dx = 0.$$

Clearly $E: y(x) \equiv 0, \lambda_0 = 1, \lambda \equiv 0$ $(0 \leqq x \leqq 2)$ is a non-singular extremal for this problem satisfying conditions II'_N and IV'_*. We shall center our attention on comparison arcs $C: Y(x)$ of the following character: $Y(2 - x) = Y(x)$ and

$$\begin{aligned} Y(x) &= m^2 x, & 0 \leqq x \leqq m^{-3}, \\ &= -m^2 x + 2m^{-1}, & m^{-3} \leqq x \leqq 2m^{-3}, \\ &\equiv 0, & 2m^{-3} \leqq x \leqq 1, \qquad (m > 2). \end{aligned}$$

Such an arc C obviously satisfies (4.11) and (4.12). The defining function $Y(x)$ is identically zero except on two intervals of total length $4/m^3$; moreover the maximum value of $Y(x)$ is $1/m$, and the length of the curve C is $2 + 4[(1 + m^4)^{\frac{1}{2}} - 1]/m^3$. For such an arc C, however, the value of

$$U(x) = \int_x^2 (1 - t)[y'(t)]^2\, dt = -\int_0^x (1 - t)[y'(t)]^2\, dt$$

is readily seen to be less than $-2(1 - 2/m^3)m$ on $2/m^3 \leqq x \leqq 2 - 2/m^3$. In particular, for preassigned quantities $\epsilon > 0$, $K > 0$, there exists a value of m such that the length of C differs from that of E by not more than ϵ, whereas, $|U(1)| > K$.

5. **Proof of Theorem** 2.2. Theorem 2.2 is a ready consequence of the results of the last section. In the first place, it is to be noted that there exists an $\eta > 0$ such that if

$$Y_i(x),\ U_s(x) = \int_x^{X_2} h_s[t, Y(t), Y'(t)]\, dt \qquad (X_1 \leqq x \leqq X_2)$$

is an admissible arc for $\mathfrak{B}$ and its elements are in $(\mathfrak{E})_\eta$, then this set is also in the neighborhood $\mathfrak{F}$ of Theorem 2.1 and its end-values are in the neighborhood $\mathfrak{M}$ of that theorem. Let ρ and ϵ denote the constants of Theorem 4.1 corresponding to this value of η, and using this value of ρ, let ϵ' denote the constant of Theorem 4.2. Finally, let $\mathcal{F}$ be a neighborhood of E in xy-space and $\mathcal{M}$ a neighborhood of the ends of E in $[x_1, y_{i1}, x_2, y_{i2}]$-space such that if C is an admissible arc for B lying in $\mathcal{F}$ and with end-points in $\mathcal{M}$, then the elements of C lie in both of the neighborhoods $(E)_\epsilon$ and $(E)_{\epsilon'}$ of E.

Now consider an admissible arc $C: Y_i(x)$ $(X_1 \leqq x \leqq X_2)$ for B that lies in $\mathcal{F}$ and has its end-points in $\mathcal{M}$. If

$$\int_{X_1}^{X_2} \mathcal{E}^*[x, Y(x), y'(x), \lambda(x); Y'(x)]\, dx \tag{5.1}$$

does not exceed ρ, then by Theorem 4.1 the corresponding arc $\mathfrak{C}$ is in $(\mathfrak{E})_\eta$, hence in the $\mathfrak{F}$ of Theorem 2.1, and has its end-points in the neighborhood $\mathfrak{M}$ of that theorem. Consequently, for such an arc C we have by Theorem 2.1 that $J[C] \geqq J[E]$, the equality holding only if $C \equiv E$. On the other hand, if C is an admissible arc for B that lies in $\mathcal{F}$, has its end-points in $\mathcal{M}$, and renders

the integral (5.1) a value greater than ρ, it follows from Theorem 4.2 that $J[C] > J[E]$.

6. **Osgood theorems.** In this section we shall prove an Osgood theorem for the problem B formulated in §2. This result is a consequence of an Osgood theorem for a general problem of Bolza not involving isoperimetric conditions, together with the Lindeberg Theorem 4.2. For a discussion of Osgood theorems for simpler problems of the calculus of variations, particularly those in parametric form, the reader is referred to Bolza [2].

THEOREM 6.1. *Suppose that* **B** *is an ordinary problem of Bolza involving expressions* (2.2), (2.3) *and* (2.5), *and that* $\mathbf{E}$: $y_i(x)$, $\lambda_0 = 1$, $\lambda_\alpha(x)$ $(x_1 \leqq x \leqq x_2)$ *satisfies with constants* e_μ *the multiplier rule,* II_N, *non-singularity, and* IV'_*. *Then there exists a neighborhood* **F** *of* **E** *in xy-space and a neighborhood* **M** *of the ends of* **E** *in* $[x_1, y_{i1}, x_2, y_{i2}]$*-space with the following property: corresponding to each neighborhood* **F**′ *of* **E** *interior to* **F**, *and each neighborhood* **M**′ *of the ends of* **E** *interior to* **M**, *there exists a constant* $r > 0$ *such that for every admissible arc* **C** *in* **F** *and with end-points in* **M**, *but which does not lie in* **F**′ *and have end-points in* **M**′, *we have* $J[\mathbf{C}] - J[\mathbf{E}] \geqq r$.

In the proof of this theorem use will be made of inequality (6.11) of Reid [10] from which the sufficiency theorem of that paper is deduced. Using the notation of that paper, we have under the hypotheses of Theorem 6.1 that there exist bounded neighborhoods **F** and **M**, together with a constant $r_0 > 0$ such that if **C**: $Y_i(x)$ $(X_1 \leqq x \leqq X_2)$ is an admissible arc for **B** lying in **F** and having end-points in **M**, then

$$(6.1)\quad J[\mathbf{C}] - J[\mathbf{E}] \geqq r_0 \left\{ \sum_{\nu=1}^{2} (X_\nu - x_\nu)^2 + \| h(X_\nu) \|^2 + \int_{X_1}^{X_2} \mathfrak{R}[\| h'(x) \|]\, dx \right\}.$$

In (6.1) we are using the convex function $\mathfrak{R}[t]$ introduced in §3 above, instead of the function $R[t]$ of Reid [10]; consequently, in (6.11) of Reid [10] the constants d_1, d_2 would now be replaced by the corresponding constants d_1, d_2 of Lemma 4.1 of the present paper. In (6.1) above we have also written $\mathfrak{R}[\| h'(x) \|]$, whereas the corresponding term in (6.11) of Reid [10] is $R[\| v(x) \|]$. In the notation of that paper, however, $v_i(x) = u_{ik}(x)h'_k(x)$ and $\| v \| \geqq m_0 \| h' \|$ on X_1X_2. Hence by property P_1 of $\mathfrak{R}[t]$, $\mathfrak{R}[\| v \|] \geqq \mathfrak{R}[m_0 \| h' \|] \geqq m_0 \min (1, m_0)\ \mathfrak{R}[\| h' \|]$. Finally, in terms of the constants appearing in equation (6.11) of Reid [10], the above constant r_0 may be defined as the least of the values $\kappa - \epsilon_0$, $\kappa m_0 - \epsilon_0(M_0 + 1 + 2d_2 + d_1)$, $[\tau - \epsilon_0 d_3(2d_2 + d_1)]m_0 \min (1, m_0)$. As pointed out at the end of §3 above, for such a problem **B** the expansion proof of Theorem 6.1 of Reid [10] is simplified if one uses a suitably chosen corresponding problem **B***; if such a problem **B*** is used, the constants appearing in (6.11) of that paper are somewhat simplified. Since these simplifications are not of fundamental importance, however, we shall not trouble to present them here.

Now suppose that **F** and **M** are such that for an admissible arc **C** in **F** and

with end-points in **M** inequality (6.1) holds. Suppose, moreover, that **F**′ is a given neighborhood of **E** interior to **F** and **M**′ is a given neighborhood of the ends of **E** which is interior to **M**. Then there exists a $\delta > 0$ such that if **C**: $Y_i(x)$ $(X_1 \leqq x \leqq X_2)$ is in $(\mathbf{E})_\delta$, then **C** is in **F**′ and has end-points in **M**′. Consequently, if **C** is an admissible arc for **B** in **F**, and with end-points in **M**, but which does not lie in **F**′ and have end-points in **M**′, then **C** is not in $(\mathbf{E})_\delta$; that is, either $(X_1 - x_1)^2 + (X_2 - x_2)^2 \geqq \delta^2$ or there is a value ξ on X_1X_2 such that $\| Y(\xi) - y(\xi) \| \geqq \delta$. If for such an arc **C** we have $(X_1 - x_1)^2 + (X_2 - x_2)^2 \geqq \delta^2$, it follows from (6.1) that $J[\mathbf{C}] - J[\mathbf{E}] \geqq r_0\delta^2$. Suppose, on the other hand, that there is a point ξ on X_1X_2 such that $\| Y(\xi) - y(\xi) \| \geqq \delta$. Since on X_1X_2, $Y_i(x) - y_i(x) = u_{ik}(x)h_k(x)$, and $\| Y(x) - y(x) \| \leqq M_0 \| h(x) \|$, it then follows that $\| h(\xi) \| \geqq \delta/M_0$. Now $\| h(\xi) \| \leqq \| h(\xi) - h(X_1) \| + \| h(X_1) \|$, and in view of the properties (3.5) and $\mathbf{P}_2$ of $\mathfrak{R}[t]$ we have

$$\mathfrak{R}[\| h(\xi) \|] \leqq 2\mathfrak{R}[\| h(\xi) - h(X_1) \|] + 2 \| h(X_1) \|^2. \tag{6.2}$$

Moreover,

$$h(\xi) - h(X_1) = \int_{X_1}^{\xi} h'(t)\, dt,$$

$$\| h(\xi) - h(X_1) \| \leqq \int_{X_1}^{\xi} \| h'(t) \| \, dt \leqq \int_{X_1}^{X_2} \| h'(t) \| dt.$$

Consequently,

$$\begin{aligned}\mathfrak{R}[\| h(\xi) - h(X_1) \|] &\leqq \mathfrak{R}\left[\int_{X_1}^{X_2} \| h'(t) \| \, dt\right] \\ &\leqq (X_2 - X_1) \max(1, X_2 - X_1)\mathfrak{R}\left[\frac{1}{X_2 - X_1}\int_{X_1}^{X_2} \| h'(t) \| \, dt\right]\end{aligned}$$

by property $\mathbf{P}_2$; hence by Jensen's inequality,

$$\mathfrak{R}[\| h(\xi) - h(X_1) \|] \leqq \max(1, X_2 - X_1) \int_{X_1}^{X_2} \mathfrak{R}[\| h'(t) \|]\, dt. \tag{6.3}$$

For such an admissible arc **C**, therefore,

$$\| h(X_1) \|^2 + \int_{X_1}^{X_2} \mathfrak{R}[\| h'(t) \|]\, dt \geqq \tfrac{1}{2} \min(1, 1/(X_2 - X_1))\, \mathfrak{R}[\| h(\xi) \|],$$

and by (6.1), $J[\mathbf{C}] - J[\mathbf{E}] \geqq \frac{1}{2}r_0 \min(1, 1/(X_2 - X_1))\, \mathfrak{R}[\delta/M_0]$. Since by hypothesis **M** is a bounded neighborhood of the ends of **E**, there is a constant $\delta_1 > 0$ such that for admissible arcs **C** with ends in **M** we have $\min(1, 1/(X_2 - X_1)) \geqq \delta_1$. Theorem 6.1 is true, therefore, for $r = \min(r_0\delta^2, (r_0\frac{1}{2}\delta_1)\, \mathfrak{R}[\delta/M_0])$.

If for an ordinary problem **B** of Bolza without isoperimetric side conditions the neighborhoods **F** and **M** of an extremal **E** and its end-points satisfy the conditions of Theorem 6.1, these neighborhoods will, for brevity, be referred to as *Osgood neighborhoods*. Since the problem $\mathfrak{B}$ of §2 is an ordinary problem of

Bolza, under the hypotheses of Theorem 2.1 there clearly exist Osgood neighborhoods of the extremal $\mathfrak{E}$ for $\mathfrak{B}$.

THEOREM 6.2. *Suppose that the hypotheses of Theorem* 2.1 *are satisfied, and that* $\mathfrak{F}$ *and* $\mathfrak{M}$ *are Osgood neighborhoods of the extremal* $\mathfrak{E}$ *for* $\mathfrak{B}$. *If* $\mathcal{F}$ *and* $\mathcal{M}$ *denote the neighborhoods for the corresponding extremal* E *for* B *as determined in the proof of Theorem* 2.2 *in* §5, *then* $\mathcal{F}$ *and* $\mathcal{M}$ *are Osgood neighborhoods for* B. *That is, corresponding to each neighborhood* $\mathcal{F}'$ *of* E *interior to* $\mathcal{F}$ *and each neighborhood* $\mathcal{M}'$ *of the end-points of* E *interior to* $\mathcal{M}$, *there exists a constant* $r' > 0$ *such that for every admissible arc* C *lying in* $\mathcal{F}$ *and with end-points in* $\mathcal{M}$, *but which does not lie in* $\mathcal{F}'$ *and have end-points in* $\mathcal{M}'$, *we have* $J[C] - J[E] \geqq r'$.

For suppose that $\mathcal{F}'$ is a neighborhood of E in xy-space which is interior to $\mathcal{F}$, and $\mathcal{M}'$ is a neighborhood of the end-points of E interior to $\mathcal{M}$. Denote by ϵ'' a positive constant such that if C: $Y_i(x)$ $(X_1 \leqq x \leqq X_2)$ is an admissible arc for B in $(E)_{\epsilon''}$, then C is in $\mathcal{F}'$ and its end-points are in $\mathcal{M}'$. Let $\mathfrak{F}'$ be a neighborhood of $\mathfrak{E}$ in xyu-space interior to $\mathfrak{F}$ and $\mathfrak{M}'$ a neighborhood of the ends of $\mathfrak{E}$ interior to $\mathfrak{M}$ such that if $\mathfrak{C}$ is an admissible arc for $\mathfrak{B}$ lying in $\mathfrak{F}'$ and having end-points in $\mathfrak{M}'$, then $\mathfrak{C}$ lies in $(\mathfrak{E})_{\epsilon''}$. As $\mathfrak{F}$ and $\mathfrak{M}$ are Osgood neighborhoods of $\mathfrak{E}$, let r^* denote the constant corresponding to $\mathfrak{F}'$ and $\mathfrak{M}'$ determined by Theorem 6.1.

Now suppose that C is an admissible arc for B lying in $\mathcal{F}$ and having end-points in $\mathcal{M}$, but which does not lie in $\mathcal{F}'$ and have its end-points in $\mathcal{M}'$. Then C does not lie in $(E)_{\epsilon''}$, the corresponding admissible arc $\mathfrak{C}$ for $\mathfrak{B}$ does not lie in $(\mathfrak{E})_{\epsilon''}$, and consequently $\mathfrak{C}$ does not lie in $\mathfrak{F}'$ and have its end-points in $\mathfrak{M}'$. If for C the expression (5.1) does not exceed ρ, it then follows as in §5 that $\mathfrak{C}$ is in $\mathfrak{F}$ and has its end-points in $\mathfrak{M}$. We then have $J[C] - J[E] \geqq r^*$. On the other hand, if the expression (5.1) exceeds ρ, it follows from Theorem 4.2 that $J[C] - J[E] \geqq \frac{1}{2}\rho$. Consequently, if C is an admissible arc for B lying in $\mathcal{F}$ and having its end-points in $\mathcal{M}$, but which does not lie in $\mathcal{F}'$ and have its end-points in $\mathcal{M}'$, we have $J[C] - J[E] \geqq r'$, where $r' = \min(r^*, \frac{1}{2}\rho)$.

BIBLIOGRAPHY

1. G. A. BLISS, *The problem of Bolza in the calculus of variations*, Mimeographed notes of lectures delivered at the University of Chicago, Winter Quarter of 1935.
2. OSCAR BOLZA, *Vorlesungen über Variationsrechnung*, Teubner, 1909.
3. M. R. HESTENES, *Sufficient conditions for the problem of Bolza in the calculus of variations*, Transactions of the American Mathematical Society, vol. 36(1934), pp. 793–818.
4. M. R. HESTENES, *A direct sufficiency proof for the problem of Bolza in the calculus of variations*, Transactions of the American Mathematical Society, vol. 42(1937), pp. 141–154.
5. M. R. HESTENES, *Generalized problems of Bolza in the calculus of variations*, this Journal, vol. 5(1939), pp. 309–324.
6. M. R. HESTENES AND W. T. REID, *A note on the Weierstrass condition in the calculus of variations*, Bulletin of the American Mathematical Society, vol. 45(1939), pp. 471–473.

7. E. E. Levi, *Sopra un teorema del calcolo delle variazioni del Sig. Lindeberg*, Rendiconti del Circolo Matematico di Palermo, vol. 37(1914), pp. 245–248.

8. J. W. Lindeberg, *Über einige Fragen der Variationsrechnung*, Mathematische Annalen, vol. 67(1909), pp. 340–354.

9. I. E. Perlin, *Sufficient conditions for a minimum in the problem of Lagrange with isoperimetric conditions*, Dissertation [Chicago, 1935], published in Contributions to the Calculus of Variations, 1933–37, The University of Chicago.

10. W. T. Reid, *Sufficient conditions by expansion methods for the problem of Bolza in the calculus of variations*, Annals of Mathematics, vol. 38(1937), pp. 662–678.

11. W. T. Reid, *A direct expansion proof of sufficient conditions for the non-parametric problem of Bolza*, Transactions of the American Mathematical Society, vol. 42(1937), pp. 183–190.

12. W. T. Reid, *A theorem on quadratic forms*, Bulletin of the American Mathematical Society, vol. 44(1938), pp. 437–440.

13. Leonida Tonelli, *Su due proposizioni di J. W. Lindeberg e E. E. Levi nel calcolo delle variazioni*, Atti della R. Accademia dei Lincei, Rendiconti, Serie V, vol. 30_1(1921), pp. 110–112.

14. Leonida Tonelli, *Fondamenti di Calcolo delle Variazioni*, vols. 1 and 2, Bologna, 1921.

University of Chicago.

THE CURVES OF A CONJUGATE NET

By E. P. Lane and M. L. MacQueen

1. **Introduction.** The purpose of this paper is to make some contributions to the projective differential geometry of the curves of a conjugate net on an analytic surface in ordinary space. §2 contains a summary of portions of the theory of a surface referred to a conjugate net. Power series expansions in non-homogeneous projective coördinates for the parametric curves on the surface are then computed to terms of the sixth degree. Some geometrical applications of these series are next made in a discussion of quadric surfaces having contact of the second order at a point of the surface. In the last two sections conjugate nets with specialized families are considered.

2. **Analytic basis.** In this section we indicate an analytic basis for the study of a surface referred to a conjugate net.

If the projective homogeneous coördinates $x^{(1)}, \cdots, x^{(4)}$ of a point P_x in ordinary space are given as analytic functions of two independent variables u, v by equations of the form

$$x = x(u, v), \tag{1}$$

the locus of P_x as u, v vary is an analytic surface S. If the parametric curves on the surface form a conjugate net, the four coördinates x and the four coördinates y of a point on the axis of the point P_x satisfy a completely integrable system of partial differential equations of the form[1]

$$\begin{aligned} x_{uu} &= px + \alpha x_u + Ly, \\ x_{uv} &= cx + ax_u + bx_v, \\ x_{vv} &= qx + \delta x_v + Ny \qquad (LN \neq 0). \end{aligned} \tag{2}$$

Let the point P_y be the harmonic conjugate of the point P_x with respect to the two foci of the axis. It is easy to verify that

$$y_u = fx - nx_u + sx_v + Ay, \qquad y_v = gx + tx_u + nx_v + By, \tag{3}$$

where we have placed

$$\begin{aligned} fN &= c_v + ac + bq - c\delta - q_u, & gL &= c_u + bc + ap - c\alpha - p_v, \\ -nN &= a_v + a^2 - a\delta - q, & tL &= a_u + ab + c - \alpha_v, \\ sN &= b_v + ab + c - \delta_u, & nL &= b_u + b^2 - b\alpha - p, \\ A &= b - (\log N)_u, & B &= a - (\log L)_v. \end{aligned} \tag{4}$$

Received May 16, 1939.

[1] E. P. Lane, *Projective Differential Geometry of Curves and Surfaces*, Chicago, 1932, p. 138.

The ray-points, or Laplace transformed points x_{-1}, x_1, of the point P_x are given by the formulas

$$x_{-1} = x_u - bx, \qquad x_1 = x_v - ax. \tag{5}$$

The following formulas give some of the invariants of the parametric conjugate net:

$$\begin{aligned} H &= c + ab - a_u, & K &= c + ab - b_v, \\ \mathfrak{H} &= sN, & \mathfrak{K} &= tL, \\ 8\mathfrak{B}' &= 4a - 2\delta + (\log r)_v, & 8\mathfrak{C}' &= 4b - 2\alpha - (\log r)_u, \\ \mathfrak{D} &= -2nL, & r &= N/L. \end{aligned} \tag{6}$$

We shall employ the covariant tetrahedron x, x_{-1}, x_1, y as a local tetrahedron of reference with a unit point chosen so that a point

$$X = y_1 x + y_2 x_{-1} + y_3 x_1 + y_4 y \tag{7}$$

has local coördinates proportional to $y_1, \cdots, y_4$. In this coördinate system the equations of the osculating planes of the parametric curves C_u, C_v at the point P_x are respectively $y_3 = 0$, $y_2 = 0$.

3. **Power series expansions.** An analytic curve in ordinary space can be defined by expressing two of the non-homogeneous coördinates of a point on the curve as power series in the third coördinate. Such power series expansions for the u-curve at a point P_x of the surface (1) may be calculated in the following way.

The coördinates X of a point near P_x and on the u-curve through P_x are expressible by Taylor's expansion as power series in the increment Δu corresponding to displacement from P_x to the point X along the u-curve,

$$X = x + x_u \Delta u + \tfrac{1}{2} x_{uu} \Delta u^2 + \tfrac{1}{6} x_{uuu} \Delta u^3 + \tfrac{1}{24} x_{uuuu} \Delta u^4 + \cdots. \tag{8}$$

Expressing each of x_{uu}, x_{uuu}, $\cdots$ as a linear combination of x, x_{-1}, x_1, y, we find that X can be expressed in the form (7), where the local coördinates $y_1, \cdots, y_4$ of the point X are given by the expansions

$$\begin{aligned} y_1 &= 1 + b\Delta u + \tfrac{1}{2}(p + b\alpha)\Delta u^2 + \tfrac{1}{6}A_3 \Delta u^3 + \tfrac{1}{24}A_4 \Delta u^4 + \cdots, \\ y_2 &= \Delta u + \tfrac{1}{2}\alpha \Delta u^2 + \tfrac{1}{6}(\alpha_u + \alpha^2 + p - nL)\Delta u^3 + \tfrac{1}{24}B_4 \Delta u^4 \\ &\qquad + \tfrac{1}{120}B_5 \Delta u^5 + \cdots, \\ y_3 &= \frac{1}{6}\frac{\mathfrak{H}}{r}\Delta u^3 + \frac{1}{24}\frac{\mathfrak{H}}{r}[\alpha + 2b - 2l_u + (\log \mathfrak{H})_u]\Delta u^4 \\ &\qquad + \frac{1}{120}\frac{\mathfrak{H}}{r}C_5 \Delta u^5 + \frac{1}{720}\frac{\mathfrak{H}}{r}C_6 \Delta u^6 + \cdots, \\ y_4 &= \tfrac{1}{2}L\Delta u^2 + \tfrac{1}{6}L(\alpha + b - l_u)\Delta u^3 + \tfrac{1}{24}LD_4 \Delta u^4 + \tfrac{1}{120}LD_5 \Delta u^5 \\ &\qquad + \tfrac{1}{720}LD_6 \Delta u^6 + \cdots, \end{aligned} \tag{9}$$

wherein l is defined by

$$l = \log r$$

and the coefficients $A_3, \cdots, D_6$ are defined by the following formulas:

$$\begin{aligned}
A_3 &= (p_u + \alpha p + fL) + b(\alpha_u + \alpha^2 + p - nL) + asL,\\
A_4 &= (p_u + \alpha p + fL)_u + b(p_u + \alpha p + fL) + (p + b\alpha)(\alpha_u + \alpha^2 + p - nL)\\
&\quad + b(\alpha_u + \alpha^2 + p - nL)_u + sL(2ab + c) + a(sL)_u\\
&\quad + (f + as - bn)(L_u + \alpha L + AL),\\
B_4 &= (p_u + \alpha p + fL) + \alpha(\alpha_u + \alpha^2 + p - nL) + (\alpha_u + \alpha^2 + p - nL)_u\\
&\quad + asL - n(L_u + \alpha L + AL),\\
B_5 &= \alpha(p_u + \alpha p + fL) + 2(p_u + \alpha p + fL)_u + (\alpha_u + \alpha^2 + p - nL)^2\\
&\quad + 2\alpha(\alpha_u + \alpha^2 + p - nL)_u + (\alpha_u + \alpha^2 + p - nL)_{uu} + 2a(sL)_u\\
&\quad + sL(c + ab + a_u + a\alpha) + (f + as - nA - \alpha n - n_u)(L_u + \alpha L + AL)\\
&\quad - 2n(L_u + \alpha L + AL)_u\,,\\
C_5 &= (\alpha_u + \alpha^2 + p - nL) + b^2 + b_u + 2b(sL)_u/sL + (sL)_{uu}/sL\\
&\quad + (\alpha + b - l_u)(A + b + s_u/s) + 2(L_u + \alpha L + AL)_u/L,\\
C_6 &= (p_u + \alpha p + fL) + (\alpha_u + \alpha^2 + p - nL)(\alpha + b + A + 2L_u/L + s_u/s)\\
&\quad + 3(\alpha_u + \alpha^2 + p - nL) + asL + b^3 + 3bb_u + b_{uu} + 3(b_u + b^2)(sL)_u/sL\\
&\quad + 3b(sL)_{uu}/sL + (sL)_{uuu}/sL + (\alpha + b - l_u)(b_u + b^2 + 2bs_u/s\\
&\quad + s_{uu}/s + bA + 2A_u + As_u/s + A^2 - nL)\\
&\quad + 3(L_u + \alpha L + AL)_u(b + A + s_u/s)/L + 3(L_u + \alpha L + AL)_{uu}/L,\\
D_4 &= (\alpha_u + \alpha^2 + p - nL) + A(\alpha + b - l_u) + (L_u + \alpha L + AL)_u/L,\\
D_5 &= (p_u + \alpha p + fL) + (\alpha_u + \alpha^2 + p - nL)(\alpha + b - l_u) + asL\\
&\quad + 2(\alpha_u + \alpha^2 + p - nL)_u + 2A(L_u + \alpha L + AL_u/L\\
&\quad + (\alpha + b - l_u)(A^2 + A_u - nL) + (L_u + \alpha L + AL)_{uu}/L,\\
D_6 &= (p_u + \alpha p + fL)(\alpha + b - l_u) + 3(p_u + \alpha p + fL)_u\\
&\quad + (\alpha_u + \alpha^2 + p - nL)(\alpha^2 + p + \alpha A + \alpha L_u/L + 2\alpha_u + A^2 + A_u\\
&\quad + 2AL_u/L + L_{uu}/L - nL) + 3(\alpha + b - l_u)(\alpha_u + \alpha^2 + p - nL)_u\\
&\quad + 3(\alpha_u + \alpha^2 + p - nL)_{uu} + sL(c + a\alpha + aA + aL_u/L + 2a_u + b)\\
&\quad + 3a(sL)_u + (\alpha + b - l_u)(fL - \alpha nL - 2nAL - 2n_uL - nL_u + asL\\
&\quad + A^3 + 3AA_u + A_{uu}) + 3(L_u + \alpha L + AL)_u(A^2 + A_u - nL)/L\\
&\quad + 3A(L_u + \alpha L + AL)_{uu}/L + (L_u + \alpha L + AL)_{uuu}/L.
\end{aligned} \tag{10}$$

Introducing non-homogeneous coördinates by the definitions

$$x = y_2/y_1\,, \qquad y = y_3/y_1\,, \qquad z = y_4/y_1\,, \tag{11}$$

we find, by use of (9), the following expansions:

$$\begin{aligned} x = {} & \Delta u + \tfrac{1}{2}(\alpha - 2b)\Delta u^2 + \tfrac{1}{6}[(\alpha - 2b)_u + (\alpha - 2b)^2 + nL]\Delta u^3 \\ & + \tfrac{1}{24}[B_4 - 4A_3 - 6(\alpha - 2b)(p + b\alpha) - 4b(\alpha - 2b)_u - 4b(\alpha - 2b)^2 \\ & - 4bnL]\Delta u^4 + \tfrac{1}{120}\{B_5 - 5A_4 - 10(\alpha - 2b)A_3 - 10(p + b\alpha)[(\alpha - 2b)_u \\ & + (\alpha - 2b)^2 + nL] - 5b[B_4 - 4A_3 - 6(p + b\alpha)(\alpha - 2b) - 4b(\alpha - 2b)_u \\ & \qquad - 4b(\alpha - 2b)^2 - 4bnL]\}\Delta u^5 + \cdots, \\ y = {} & \frac{1}{6}\frac{\mathfrak{H}}{r}\Delta u^3 + \frac{1}{24}\frac{\mathfrak{H}}{r}[\alpha - 2b - 2l_u + (\log \mathfrak{H})_u]\Delta u^4 \\ & + \frac{1}{120}\frac{\mathfrak{H}}{r}\{C_5 - 10(p + b\alpha) - 5b[\alpha - 2b - 2l_u + (\log \mathfrak{H})_u]\}\Delta u^5 \\ & + \frac{1}{720}\frac{\mathfrak{H}}{r}\{C_6 - 6bC_5 - 20A_3 + 15(2b^2 - p - b\alpha)[\alpha - 2b - 2l_u \\ & \qquad + (\log \mathfrak{H})_u] + 60(p + b\alpha)\}\Delta u^6 + \cdots, \\ z = {} & \tfrac{1}{2}L\Delta u^2 + \tfrac{1}{6}L(\alpha - 2b - l_u)\Delta u^3 \\ & + \tfrac{1}{24}L[D_4 - 6(p + b\alpha) - 4b(\alpha - 2b - l_u)]\Delta u^4 \\ & + \tfrac{1}{120}L[D_5 - 5bD_4 - 10A_3 + 10(\alpha - 2b - l_u)(2b^2 - p - b\alpha) \\ & + 30b(p + b\alpha)]\Delta u^5 + \tfrac{1}{720}L[D_6 - 6bD_5 - 15(p + b\alpha)D_4 \\ & + 30b^2D_4 - 15A_4 + 60bA_3 - 20(\alpha - 2b - l_u)A_3 \\ & - 120b(\alpha - 2b - l_u)(b^2 - p - b\alpha) + 90(p + b\alpha)^2 \\ & \qquad - 180b^2(p + b\alpha)]\Delta u^6 + \cdots. \end{aligned} \tag{12}$$

Writing y and z as power series in x with undetermined coefficients and demanding that these series be satisfied by the series (12) identically in Δu as far as the terms of the sixth order, we obtain *the following power series expansions for the u-curve at a point P_x of a conjugate net*:

$$\begin{aligned} y &= \frac{1}{6}\frac{\mathfrak{H}}{r}x^3 + \frac{1}{24}\frac{\mathfrak{H}}{r}(I + 16\mathfrak{C}')x^4 + \frac{1}{120}\frac{\mathfrak{H}}{r}a_5x^5 + \frac{1}{720}\frac{\mathfrak{H}}{r}a_6x^6 + \cdots, \\ z &= \tfrac{1}{2}Lx^2 + \tfrac{4}{3}L\mathfrak{C}'x^3 + \tfrac{1}{24}Lb_4x^4 + \tfrac{1}{120}Lb_5x^5 + \tfrac{1}{720}Lb_6x^6 + \cdots, \end{aligned} \tag{13}$$

where the coefficients $a_5\,, \cdots, b_6$ are defined by the formulas

$$\begin{aligned} b_4 &= 4(2\mathfrak{C}'_u + \mathfrak{C}'l_u + 24\mathfrak{C}'^2), \\ a_5 &= 3b_4 + 24\mathfrak{C}'I + \tfrac{3}{2}I^2 + J, \\ b_5 &= b_{4u} + (16\mathfrak{C}' + l_u)b_4 - 8\mathfrak{C}'\mathfrak{D} + 6LP, \\ a_6 &= 4b_5 + 6Ib_4 + E + J_u + Jl_u - \tfrac{1}{2}I\mathfrak{D} + 2LP, \\ b_6 &= b_{5u} + (60\mathfrak{C}' + \tfrac{3}{2}l_u)b_5 - 8\mathfrak{C}'(188\mathfrak{C}' + 5l_u)b_4 - 40\mathfrak{C}'b_{4u} + 10\,\frac{H\mathfrak{H}}{r}, \end{aligned} \tag{14}$$

in which I, J, E, and P are defined by placing

$$
\begin{aligned}
I &= (\log \mathfrak{H})_u + 4\mathfrak{C}' + \tfrac{1}{2}l_u\,, \\
J &= I_u - \tfrac{1}{2}I^2 + \tfrac{1}{2}Il_u + 4\mathfrak{C}'I + \mathfrak{D}, \\
E &= 3I^3 + 4IJ + 40\mathfrak{C}'J + 48\mathfrak{C}'I^2, \\
P &= f + as - bn.
\end{aligned}
\tag{15}
$$

Analogous expansions for the v-curve at the point P_x can be written by making the appropriate symmetrical interchanges of the symbols.

4. **A canonical form for the expansions.** In this section it will be shown that by suitable choice of the coördinate system the power series expansions (13) can be reduced to an especially simple canonical form.

The parametric equations of the osculating twisted cubic at the point P_x of the u-curve can be written in the form

$$
\begin{aligned}
y_1 &= 1 + \tfrac{1}{2}It + Ct^2 + Dt^3, \\
y_2 &= t + Gt^2 + Ft^3, \\
y_3 &= \frac{1}{6}\frac{\mathfrak{H}}{r}t^3, \\
y_4 &= \tfrac{1}{2}Lt^2,
\end{aligned}
\tag{16}
$$

where we have placed

$$
\begin{aligned}
G &= \tfrac{1}{4}I - \tfrac{4}{3}\mathfrak{C}', \\
F &= \tfrac{1}{20}(J + \tfrac{1}{4}I^2 - \tfrac{1}{3}b_4 - \tfrac{8}{3}\mathfrak{C}'I + \tfrac{320}{9}\mathfrak{C}'^2), \\
D &= \tfrac{1}{60}b_5 - \tfrac{2}{3}\mathfrak{C}'b_4 + \tfrac{1}{2}IF + \tfrac{1024}{27}\mathfrak{C}'^3, \\
C &= \tfrac{1}{10}(J + \tfrac{7}{8}I^2 + \tfrac{1}{2}b_4 - \tfrac{8}{3}\mathfrak{C}'I - \tfrac{160}{3}\mathfrak{C}'^2).
\end{aligned}
\tag{17}
$$

The osculating twisted cubic (16) may also be represented by means of two power series expansions which must agree with the series (13) up to and including the terms in x^5. For this result we find

$$
\begin{aligned}
y &= \frac{1}{6}\frac{\mathfrak{H}}{r}x^3 + \frac{1}{24}\frac{\mathfrak{H}}{r}(I + 16\mathfrak{C}')x^4 + \frac{1}{120}\frac{\mathfrak{H}}{r}a_5x^5 + \frac{1}{720}\frac{\mathfrak{H}}{r}\bar{a}_6x^6 + \cdots, \\
z &= \tfrac{1}{2}Lx^2 + \tfrac{4}{3}L\mathfrak{C}'x^3 + \tfrac{1}{24}Lb_4x^4 + \tfrac{1}{120}Lb_5x^5 + \tfrac{1}{720}L\bar{b}_6x^6 + \cdots,
\end{aligned}
\tag{18}
$$

where

$$
\begin{aligned}
\bar{a}_6 &= 4b_5 + 6Ib_4 + E + \tfrac{1}{2}IJ, \\
\bar{b}_6 &= (\tfrac{3}{2}I + 56\mathfrak{C}')b_5 + \tfrac{49}{10}b_4^2 + \tfrac{1}{20}[3(4J + I^2) - 1232\mathfrak{C}'I - 35{,}840\mathfrak{C}'^2]b_4 \\
&\quad - \tfrac{9}{160}(4J + I^2)^2 + \tfrac{2}{5}\mathfrak{C}'(4J + I^2)(3I - 40\mathfrak{C}') \\
&\quad - \tfrac{32}{5}\mathfrak{C}'^2(I^2 - 560\mathfrak{C}'I - 11{,}200\,\mathfrak{C}'^2).
\end{aligned}
\tag{19}
$$

Let us make the transformation

$$
(20) \qquad
\begin{aligned}
x &= \frac{X + FY + GZ}{1 + \frac{1}{2}IX + DY + CZ}, \\
y &= \frac{\frac{1}{6}(\mathfrak{H}/r)Y}{1 + \frac{1}{2}IX + DY + CZ}, \\
z &= \frac{\frac{1}{2}LZ}{1 + \frac{1}{2}IX + DY + CZ}
\end{aligned}
$$

from the old coördinates x, y, z to new coördinates X, Y, Z. This transformation moves the vertex $(0, 0, 0, 1)$ of the local tetrahedron of reference to a point on the osculating cubic of the u-curve. Moreover, the edge $X_1 = X_2 = 0$ of the new tetrahedron is tangent to the osculating cubic at the point $(0, 0, 0, 1)$, and the face $X_1 = 0$ is the osculating plane of the cubic at this point. The unit point is on the osculating cubic whose equations become

$$
(21) \qquad Y = X^3, \qquad Z = X^2.
$$

Thus we find that *the power series expansions for the u-curve at the point P_x can be written in the canonical form*

$$
(22) \qquad
\begin{aligned}
Y &= X^3 + A_6X^6 + \cdots, \\
Z &= X^2 + B_6X^6 + \cdots,
\end{aligned}
$$

in which the coefficients A_6, B_6 are defined by the formulas

$$
(23) \qquad
\begin{aligned}
A_6 &= \tfrac{1}{120}(a_6 - \bar{a}_6), \\
B_6 &= \tfrac{1}{360}(b_6 - \bar{b}_6).
\end{aligned}
$$

5. **Applications to quadrics having second-order contact with the surface.** The power series expansions deduced in the preceding sections will now be employed to investigate certain configurations which are covariantly associated with a point of a conjugate net of a surface.

First of all, the equation of any quadric surface having contact of the second order with the surface at the point P_x can be written in the form

$$
(24) \qquad Lx^2 + Ny^2 - 2z + k_2xz + k_3yz + k_4z^2 = 0,
$$

where k_2, k_3, k_4 are arbitrary. By means of equations (13) and the analogous expansions for the v-curve, it is easy to show that the quadrics (24) for which

$$
(25) \qquad k_2 = \tfrac{16}{3}\mathfrak{C}', \qquad k_3 = \tfrac{16}{3}\mathfrak{B}'
$$

are the quadrics having contact of the third order with both the u-curve and the v-curve at the point P_x. Thus we find that *any quadric of the pencil*

$$
(26) \qquad Lx^2 + Ny^2 - 2z + \tfrac{16}{3}\mathfrak{C}'xz + \tfrac{16}{3}\mathfrak{B}'yz + k_4z^2 = 0,
$$

where k_4 is arbitrary, has contact of the third order with the parametric curves at a point P_x of the surface. If a unique quadric of this pencil is desired, we may

choose the one that passes through the covariant point P_y. For this quadric we have $k_4 = 0$. Incidentally, it may be remarked that the unique quadric for which $k_2 = k_3 = k_4 = 0$ in equation (24) has been called[2] the canonical quadric of the parametric conjugate net. The axis and ray are polar reciprocal lines with respect to this quadric which passes through the point P_y.

Among the quadrics (24) there is a pencil having contact of the fourth order with the u-curve (13) at the point P_x. For this pencil we find

$$k_2 = \frac{16}{3}\mathfrak{C}', \qquad k_4 = \frac{4}{9L}(6\mathfrak{C}'_u + 3\mathfrak{C}'l_u + 8\mathfrak{C}'^2), \tag{27}$$

with k_3 arbitrary. It is of interest to observe that the quadric of Moutard for the u-tangent[3] is a unique quadric of this pencil, and for this quadric we have $k_3 = 0$. Finally, there is a unique quadric of this pencil having fifth-order contact with the u-curve at the point P_x, and for this quadric we find

$$\frac{\mathfrak{H}}{r}k_3 = \frac{1}{5}b_5 - 8\mathfrak{C}'b_4 + \frac{4096}{9}\mathfrak{C}'^3. \tag{28}$$

The equations of the analogous quadrics associated with the v-curve at the point P_x can be written without difficulty.

Lane has shown[4] that the coördinates of the principal points of Bompiani[5] of the two curves of the parametric conjugate net at a point P_x of a surface are

$$x_{-1} + \tfrac{8}{3}\mathfrak{C}'x \pm (x_1 + \tfrac{8}{3}\mathfrak{B}'x)r^{-\frac{1}{2}}. \tag{29}$$

These points are evidently on the associate conjugate tangents

$$Lx^2 - Ny^2 = 0, \qquad z = 0, \tag{30}$$

and on the principal join whose equations are

$$8(\mathfrak{C}'x + \mathfrak{B}'y) = 3, \qquad z = 0. \tag{31}$$

A geometric characterization of the principal points of the parametric curves at the point P_x is contained in the following theorem, the truth of which is easy to verify.

At a point P_x of a conjugate net the principal point on each associate conjugate tangent is the pole of the plane determined by the axis and the other associate conjugate tangent with respect to any quadric of the pencil (26).

Moreover, it is then obvious that *the principal join of the fundamental para-*

[2] W. M. Davis, *Contributions to the theory of conjugate nets*, Chicago doctoral dissertation (1932), p. 10.

[3] M. L. MacQueen, *On the principal join of two curves on a surface*, American Journal of Mathematics, vol. 58(1936), p. 624.

[4] E. P. Lane, *Invariants of intersection of two curves on a surface*, American Journal of Mathematics, vol. 54(1932), pp. 699–706.

[5] E. Bompiani, *Invarianti d'intersezione di due curve sghembe*, Rendiconti dei Lincei, (6), vol. 14(1931), pp. 456–461.

metric curves at a point P_x of a surface is the polar line of the axis with respect to any quadric of the pencil (26).

It is easy to verify that the polar line of the ray with respect to any quadric of the pencil (26) joins the point P_x to the point

$$-\frac{8\mathfrak{C}'x_{-1}}{3L} - \frac{8\mathfrak{B}'x_1}{3N} + y, \tag{32}$$

and thus lies in the canonical plane[6] whose equation is

$$\mathfrak{B}'x - r\mathfrak{C}'y = 0. \tag{33}$$

Furthermore, the polar line of the associate axis with respect to any quadric of the pencil (26) is found to have the equations

$$2(\mathfrak{C}'x + \mathfrak{B}'y) = 3, \qquad z = 0, \tag{34}$$

and so this line passes through the canonical point

$$\mathfrak{B}'x_{-1} - \mathfrak{C}'x_1 \tag{35}$$

of Davis.

6. **Conditions for twisted cubics.** Consideration of the canonical expansions (22) leads to some interesting results. In the first place, application of the projective differential theory of curves in ordinary space shows that the u-curves belong to linear complexes in case $A_6 = 0$. By means of equations (14), (19), this condition can be reduced to

$$J_u + Jl_u - \tfrac{1}{2}IJ - \tfrac{1}{2}\mathfrak{D}I + 2LP = 0. \tag{36}$$

Thus we easily have the theorem essentially the same as a theorem formerly[7] obtained by Lane using a more laborious process:

A necessary and sufficient condition that the u-curves of a conjugate net in ordinary space belong to linear complexes is given by equation (36). Further application of the theory of curves shows that the u-curves are twisted cubics in case $A_6 = B_6 = 0$. The condition $B_6 = 0$ can be written

$$\begin{aligned} &b_{5u} - \tfrac{3}{2}(I - l_u + 24\mathfrak{C}')b_5 - \tfrac{1}{20}[3(4J + I^2) + 784\mathfrak{C}'_u + 392\mathfrak{C}'l_u - 1232\mathfrak{C}'I \\ &- 9152\mathfrak{C}'^2]b_4 + \tfrac{9}{160}(4J + I^2)^2 - \tfrac{3}{5}\mathfrak{C}'(4J + I^2)(3I - 40\mathfrak{C}') \\ &+ \tfrac{32}{5}\mathfrak{C}'^2(I^2 - 560\mathfrak{C}'I - 11{,}200\mathfrak{C}'^2) - 320\mathfrak{C}'^2\mathfrak{D} + 240\mathfrak{C}'LP + 10\frac{H\mathfrak{H}}{r} = 0. \end{aligned} \tag{37}$$

Consequently the following theorem is established:

Necessary and sufficient conditions that the u-curves of a conjugate net be twisted cubics are given by equations (36), (37).

[6] W. M. Davis, loc. cit., p. 17.

[7] E. P. Lane, *Contributions to the theory of conjugate nets*, American Journal of Mathematics, vol. 49(1927), p. 574.

7. **Segre-Darboux net.** The theory of the preceding sections will now be applied to some special cases which will be considered in this and the following section.

It is known that the u-curves of a conjugate net are curves of Darboux and the v-curves are the corresponding curves of Segre in case $\mathfrak{C}' = 0$. In this section such a net is called a Segre-Darboux net. *The expansions* (13), *when specialized for the u-curve at a point P_x of a Segre-Darboux net, become*

$$
\begin{aligned}
y &= \frac{1}{6}\frac{\mathfrak{H}}{r}x^3 + \frac{1}{24}\frac{\mathfrak{H}}{r}Ix^4 + \frac{1}{120}\frac{\mathfrak{H}}{r}\left(J + \frac{3}{2}I^2\right)x^5 + \frac{1}{720}\frac{\mathfrak{H}}{r}a_6x^6 + \cdots, \\
z &= \tfrac{1}{2}Lx^2 + \tfrac{1}{20}L^2Px^5 + \tfrac{1}{720}Lb_6x^6 + \cdots,
\end{aligned} \tag{38}
$$

where

$$
\begin{aligned}
a_6 &= J_u + Jl_u + 3I^3 + 4IJ - \tfrac{1}{2}\mathfrak{D}I + 26LP, \\
b_6 &= 9LPl_u + 6(LP)_u + 10\,\frac{H\mathfrak{H}}{r},
\end{aligned} \tag{39}
$$

and the definitions of I and J reduce to

$$
I = (\log \mathfrak{H})_u + \tfrac{1}{2}l_u\,, \qquad J = I_u + \tfrac{1}{2}Il_u - \tfrac{1}{2}I^2 + \mathfrak{D}. \tag{40}
$$

Moreover, the expansions (18) for the osculating cubic at the point P_x of the u-curve become

$$
\begin{aligned}
y &= \frac{1}{6}\frac{\mathfrak{H}}{r}x^3 + \frac{1}{24}\frac{\mathfrak{H}}{r}Ix^4 + \frac{1}{120}\frac{\mathfrak{H}}{r}\left(J + \frac{3}{2}I^2\right)x^5 + \frac{1}{720}\frac{\mathfrak{H}}{r}\bar{a}_6x^6 + \cdots, \\
z &= \tfrac{1}{2}Lx^2 + \tfrac{1}{20}L^2Px^5 + \tfrac{1}{720}L\bar{b}_6x^6 + \cdots,
\end{aligned} \tag{41}
$$

wherein

$$
\begin{aligned}
\bar{a}_6 &= 3I^3 + \tfrac{9}{2}IJ + 24LP, \\
\bar{b}_6 &= 9LPI - \tfrac{9}{160}(4J + I^2)^2.
\end{aligned} \tag{42}
$$

Parametric equations of the osculating cubic (41) are given by

$$
\begin{aligned}
y_1 &= 1 + \tfrac{1}{2}It + \tfrac{1}{80}(8J + 7I^2)t^2 + \tfrac{1}{160}(I^3 + 4IJ + 16LP)t^3, \\
y_2 &= t + \tfrac{1}{4}It^2 + \tfrac{1}{80}(4J + I^2)t^3, \\
y_3 &= \frac{1}{6}\frac{\mathfrak{H}}{r}t^3, \\
y_4 &= \tfrac{1}{2}Lt^2.
\end{aligned} \tag{43}
$$

It is known that through a non-asymptotic tangent at a point of a surface there are ordinarily two planes which produce sextactic sections of the surface. Performing the requisite calculations, we find that the equations of the two aforementioned planes which pass through a non-specialized u-tangent at a point of the surface are given by

$$
y = \rho_i z \qquad (i = 1, 2), \tag{44}
$$

where ρ_1, ρ_2 are the roots of the equation

$$LN\mathfrak{C}'\rho^2 - \tfrac{1}{4}L(H - \mathfrak{H})\rho - \tfrac{1}{40}b_5 + \mathfrak{C}'b_4 - \tfrac{5}{12}\mathfrak{C}'^2 = 0. \tag{45}$$

For the u-tangent at a point P_x of a curve of a Segre-Darboux net, the equation of one of the planes (44) is found to be

$$5(H - \mathfrak{H})y + 3Pz = 0, \tag{46}$$

and the other plane is the tangent plane, $z = 0$, at the point P_x of the surface. If $H \neq \mathfrak{H}$, the plane (46) coincides with the osculating plane at the point P_x of the u-curve in case $P = 0$. Thus we find that *through a tangent at a point P_x of a u-curve of a parametric Segre-Darboux net the plane, except the tangent plane, intersecting the surface in a curve with a sextactic point at P_x coincides with the osculating plane of the curve at the point if, and only if, $P = 0$.*

It can be shown from the results of §5 that *the quadric of Moutard for the tangent at a point P_x of a u-curve of a parametric Segre-Darboux net coincides with the canonical quadric whose equation is*

$$Lx^2 + Ny^2 - 2z = 0. \tag{47}$$

This quadric is intersected by the osculating cubic (43) in six points, five of which coincide at the point P_x. The point of intersection distinct from the point P_x is the point for which

$$\left[5\frac{\mathfrak{H}^2}{r} + \frac{9}{320}(4J + I^2)^2\right]t = 18LP. \tag{48}$$

Thus we arrive at the following result:

At a point P_x of a u-curve of a parametric Segre-Darboux net the osculating cubic of the curve intersects the quadric of Moutard for the tangent at the point of the curve in points which coincide at P_x if, and only if, $P = 0$.

Finally, it will be observed that *the u-curves of a Segre-Darboux net belong to linear complexes in case*

$$J_u + Jl_u - \tfrac{1}{2}IJ - \tfrac{1}{2}\mathfrak{D}I + 2LP = 0, \tag{49}$$

wherein I, J are defined by equations (40). *Moreover, these curves are twisted cubics if to equation* (49) *is adjoined the condition*

$$\frac{2}{3}(LP)_u - LP(I - l_u) + \frac{1}{160}(4J + I^2)^2 + \frac{10}{9}\frac{H\mathfrak{H}}{r} = 0. \tag{50}$$

8. **Nets with one family plane curves.** In this section we again specialize the general theory by supposing that the u-curves are plane curves. It is known that the u-curves of a conjugate net in ordinary space are plane curves in case $\mathfrak{H} = 0$. Therefore *the expansions* (13), *when specialized for a plane u-curve at the point P_x, become*

$$\begin{aligned} y &= 0, \\ z &= \tfrac{1}{2}Lx^2 + \tfrac{1}{3}L\mathfrak{C}'x^3 + \tfrac{1}{24}Lb_4x^4 + \tfrac{1}{120}Lb_5x^5 + \tfrac{1}{720}Lb_6x^6 + \cdots, \end{aligned} \tag{51}$$

where b_4, b_5 are defined by equations (14) and the definition of b_6 reduces to

$$b_6 = b_{5u} + (60\mathfrak{C}' + \tfrac{3}{2}l_u)b_5 - 40\mathfrak{C}'b_{4u} - 40\mathfrak{C}'l_ub_4 - 1504\mathfrak{C}'^2b_4. \tag{52}$$

The equations of the conics having contact of the third order with the curve (51) at the point P_x are easily found to be

$$\begin{aligned} z - \tfrac{1}{2}Lx^2 - \tfrac{8}{3}\mathfrak{C}'xz + hz^2 &= 0, \\ y &= 0, \end{aligned} \tag{53}$$

where h is arbitrary. For the osculating conic the parameter h has the value

$$Lh = \tfrac{128}{9}\mathfrak{C}'^2 - \tfrac{1}{6}b_4. \tag{54}$$

Let us substitute the power series (51) in the equation of the osculating conic of the curve and demand that this equation be satisfied identically in x up to and including the terms in x^5. The curve is then hyperosculated by its osculating conic, so that if this happens at every point of the curve, then the curve is a conic. Thus we prove the theorem:

If the u-curves are plane curves, they are conics if, and only if,

$$128\mathfrak{C}'^3 - 18(2\mathfrak{C}'_{uu} + \mathfrak{C}'l_{uu} + 3\mathfrak{C}'_ul_u + \mathfrak{C}'l_u^2 - 2\mathfrak{C}'\mathfrak{D}) - 27LP = 0. \tag{55}$$

The polar line of the ray-point x_{-1} with respect to any conic of the pencil (53) has the equations

$$\begin{aligned} 3Lx + 8\mathfrak{C}'z &= 0, \\ y &= 0. \end{aligned} \tag{56}$$

This line evidently coincides with the axis at the point P_x in case $\mathfrak{C}' = 0$, so that the curve (51) is then a plane curve of Darboux. Therefore *a u-curve which is a plane curve is a curve of Darboux if, and only if, the ray-point x_{-1} is the pole of the axis with respect to any conic of the pencil having contact of the third order with the curve at the point P_x.*

It follows from equations (51) that *if the u-curves are plane curves of Darboux, the expansions for the curve at the point P_x are*

$$\begin{aligned} y &= 0, \\ z &= \tfrac{1}{2}Lx^2 + \tfrac{1}{20}L^2Px^5 + \tfrac{1}{240}L[3LPl_u + 2(LP)_u]x^6 + \cdots. \end{aligned} \tag{57}$$

It may be remarked that Čech has investigated[8] the surfaces whose curves of Darboux are plane using, however, asymptotic parameters instead of conjugate parameters.

[8] Fubini and Čech, *Geometria Proiettiva Differenziale*, Bologna, 1926, vol. 1, p. 170.

If $P \neq 0$, the equations of the eight-point nodal cubic at the point P_x of the curve (57) are found to be

$$
\begin{aligned}
&z(L^2Px - \tfrac{1}{6}b_6z) = \tfrac{1}{2}L^3Px^3 - \tfrac{1}{12}Lb_6x^2z + \tfrac{2}{3}LP^2z^3, \\
(58) \qquad &y = 0,
\end{aligned}
$$

where

$$b_6 = 3LPl_u + 2(LP)_u\,.$$

The nodal tangents at the point P_x are the u-tangent,

$$y = z = 0,$$

and the projective normal of the curve whose equations are

$$
\begin{aligned}
&6L^2Px - b_6z = 0, \\
(59) \qquad &y = 0.
\end{aligned}
$$

Thus we find that *at a point of a u-curve which is a plane curve of Darboux the projective normal of the curve coincides with the axis of the net if, and only if,*

$$3LPl_u + 2(LP)_u = 0. \tag{60}$$

It is known that the coördinates of a variable point on an analytic plane curve satisfy an ordinary differential equation of the form

$$x''' + 3p_1x'' + 3p_2x' + p_3x = 0, \tag{61}$$

where accents denote differentiation with respect to the parameter along the curve. Two invariants θ_3, θ_8 of this differential equation are given by

$$
\begin{aligned}
&\theta_3 = P_3 - \tfrac{3}{2}P_2', \\
(62) \qquad &\theta_8 = 6\theta_3\theta_3'' - 7\theta_3'^2 - 27P_2\theta_3^2,
\end{aligned}
$$

where P_2, P_3 are defined by the formulas

$$
\begin{aligned}
&P_2 = p_2 - p_1^2 - p_1', \\
(63) \qquad &P_3 = p_3 - 3p_1p_2 + 2p_1^3 - p_1''.
\end{aligned}
$$

It is furthermore known that the integral curves of equation (61) are conics in case $\theta_3 = 0$, and are coincidence curves in case $\theta_8 = 0$.

If the u-curves are plane curves of Darboux, on calculating the differential equation of the form (61) for these curves and making use of equations (62), (63), one finds that the invariants θ_3, θ_8 are given by

$$
\begin{aligned}
&\theta_3 = -LP, \\
(64) \qquad &\theta_8 = 6LP(LP)_{uu} - 7(LP)_u^2 + 9(LP)^2(l_{uu} + \tfrac{1}{4}l_u^2 + \mathfrak{D}).
\end{aligned}
$$

It follows from equation (55) or from the form of the expansions (57), as well as from the vanishing of θ_3 in the first of equations (64), that *if the u-curves are plane curves of Darboux, they are conics if, and only if,*

$$P = 0.$$

In this case the expansions for the curve at the point P_x are

$$y = 0, \qquad z = \tfrac{1}{2}Lx^2. \tag{65}$$

Finally, if $P \neq 0$, equating to zero the expression for θ_8 in the second of equations (64) yields the necessary and sufficient condition that a u-curve which is a plane curve of Darboux be a coincidence curve.

THE UNIVERSITY OF CHICAGO AND SOUTHWESTERN COLLEGE.

as
ane

ons
ane

I

Am

Cha

Con

Duk

Eco

His

The

Law

The

The

PERIODICALS PUBLISHED BY DUKE UNIVERSITY

American Literature. A quarterly journal devoted to research in American Literature, published with the coöperation of the American Literature Group of the Modern Language Association of America. Subscription, $4.00 per year. Back volumes, $5.00 each.

Character and Personality. A psychological quarterly devoted to studies of behavior and personality. Subscription, $2.00 per year. The first number was published September, 1932.

Contributions to Psychological Theory. A monograph series dealing with problems of psychological theory in the widest sense, including their relations to other fields of inquiry. The monographs appear irregularly. Subscription, $5.00 per volume of approximately 450 pages.

Duke Mathematical Journal.

Ecological Monographs. A quarterly journal devoted to the publication of original researches of ecological interest from the entire field of biological science. Subscription, $6.00 per year. The first number was published January, 1931.

Hispanic American Historical Review. A quarterly review dealing with the history of the Latin-American countries. Subscription, $4.00 per year.

The Journal of Parapsychology. A scientific quarterly dealing with telepathy, clairvoyance, and other parapsychological problems. Subscription, $3.00 per year.

Law and Contemporary Problems. A quarterly published by the School of Law, presenting in each issue a symposium on a problem of current importance having significant legal aspects. Subscription, $2.00 per year. The first number was published September, 1933.

The South Atlantic Quarterly. A magazine of modern opinion and discussion, founded in 1902. Subscription, $3.00 per year.

The Southern Association Quarterly. As official organ of the Southern Association of Colleges and Secondary Schools, it contains the proceedings of the annual meeting, together with much additional material directly related to the work of the Association. Subscription, $4.00 per year.

DUKE UNIVERSITY PRESS
DURHAM, NORTH CAROLINA

CONTENTS